الجامعة الأردنيـة
كلية الدراسات العليا

التاريـــخ: ٢١ / ٨ / ٢٠١١

نموذج رقم (١٨)
اقرار والتزام بقوانين الجامعة الأردنية وأنظمتها وتعليماتها لطلبة الماجستير

أنا الطالب: أيمان علي محمد عومان الرقم الجامعي: ٨٠٩٠٧٤٦

التخصص: أنظمة المعلومات الحاسوبية الكلية: كلية الملك عبدالله الثاني لتكنولوجيا المعلومات

اعلن بأنني قد التزمت بقوانين الجامعة الأردنية وأنظمتها وتعليماتها وقراراتها السارية المفعول المتعلقة باعداد رسائل الماجستير والدكتوراة عندما قمت شخصيا" باعداد رسالتي / اطروحتي بعنوان: ....

FAST AUTOMATED DIAGNOSIS OF MALIGNANT SKIN LESIONS

وذلك بما ينسجم مع الأمانة العلمية المتعارف عليها في كتابة الرسائل والأطاريح العلمية. كما أنني أعلن بأن رسالتي /اطروحتي هذه غير منقولة أو مستلة من رسائل أو أطاريح أو كتب أو أبحاث أو أي منشورات علمية تم نشرها أو تخزينها في أي وسيلة اعلامية، وتأسيسا" على ما تقدم فانني أتحمل المسؤولية بأنواعها كافة فيما لو تبين غير ذلك بما فيه حق مجلس العمداء في الجامعة الأردنية بالغاء قرار منحي الدرجة العلمية التي حصلت عليها وسحب شهادة التخرج مني بعد صدورها دون أن يكون لي أي حق في التظلم أو الاعتراض أو الطعن بأي صورة كانت في القرار الصادر عن مجلس العمداء بهذا الصدد.

توقيع الطالب:.. التاريخ: ٢١ / ٨ / ٢٠١١

The University of Jordan
Authorization Form

I, Aman Ali Mohammed Awman authorize the University of Jordan to supply copies of my Thesis/Dissertation to libraries or establishments or individuals on request.

Signature:
Date: 21 / 8 / 2011

FAST AUTOMATED DIAGNOSIS OF MALIGNANT SKIN LESIONS

By
Aman Ali Awman

Supervisor
Dr. Mohammad Belal R. Alzoubi, Prof.

Co-Supervisor
Dr. Bareqa I. Salah

This Thesis was Submitted in Partial Fulfillment of the Requirements for the Master's Degree of Computer Information Systems

Faculty of Graduate Studies
The University of Jordan

August, 2011

COMMITTEE DECISION

This Thesis/Dissertation (Fast Automated Diagnosis of Malignant Skin Lesions) was successfully Defended and Approved on 9/8/2011

<u>**Examination Committee**</u>	<u>**Signature**</u>
Dr. Mohammad Belal Raja Alzoubi, (Supervisor) Prof. of Data Mining	------------------------------
Dr. Bareqa Ibraheem Salah (Co-Supervisor) Assoc. Prof. of Medicine	------------------------------
Dr. Amjad Ahmad Hudaib (Member) Assoc. Prof. of Software Engineering	------------------------------
Dr. Mohammad Aref Alshraideh (Member) Assist. Prof. of Software Testing Using Evolutionary Algorithms	------------------------------
Dr. Mohammad Ali Abbadi (Member) Assoc. Prof. of Image Processing (Mutah University)	------------------------------

List of Contents

List of Tables

LIST OF FIGURES

LIST OF ABBREVIATIONS

ABCD rule: Asymmetry Boundary Color Diameter

ABCDE rule: Asymmetry Boundary Color Diameter Evolving

AI : Artificial Intelligence

ANN: Artificial Neural Network

BCC : Basal Cell Carcinoma

CAD: Computer Aided Diagnosis

FN: False Negative

FP: False Positive

GP: General Practitioner

ML: Machine Learning

NN: Neural Network

PCA: Principal Components Analysis

RGB: Red Green Blue

ROI: Region of Interest

SCC: Squamous Cell Carcinoma

SVM: Support Vector Machine

TN: True Negative

TP: True Positive

LIST OF APPENDICES

FAST AUTOMATED DIAGNOSIS OF MALIGNANT SKIN LESIONS

By
Aman A. Awman

Supervisor
Dr. Mohammed Belal R. Alzoubi, Prof.

Co-Supervisor
Dr. Bareqa I. Salah

ABSTRACT

Skin Cancer is the most common cancer all over the world. About 75% of skin cancer deaths are caused by malignant melanoma, while the survival rate for patients whose melanoma is detected early is about 99%.

Skin cancer diagnosis systems serve as diagnostic adjuncts for medical professionals for the confirmation of a diagnosis, as well as for training new dermatologists and enhancing the quality of medical care, particularly in areas where a specialized physician is not available.

This thesis proposes an automated diagnosis system of malignant skin lesions. Many researches had been done addressing image processing techniques to locate the tumor in the images and calculated relevant features that had been used to determine the type of skin cancer. Feed forward Neural Network (NN) is used to distinguish malignant melanoma types from non-melanoma types. Using Principal Components Analysis (PCA) method to decrease the dimensionality of data space is studied on assumption to enhance the fastness of classification process, where the data space had been reduced first to three dimensions (the first three principal components) then had been reduced to eight dimensions (the first eight principal components) in the second time.

The accuracy rate of NN classifier by using the 20 extracted features as inputs was 90.86% with average time 1.2051 seconds, while it was 87.59% with average time 1.1434 seconds by using the first three components, and 90.17% with average time 1.0576 seconds by using the first eight components. The sensitivity rate of NN by using the 20 extracted features as inputs was 93.36% and the specificity was 87.20%, while the sensitivity was 91.85% by using the first three components and the specificity was 80.72%, and the sensitivity was 93.71% by using the first eight components and the specificity was 85.61%.

We conclude that the NN performance indices (accuracy, sensitivity and specificity) gained using the first eight components are somehow sensible and acceptable. Using the eight components decrease the dimensionality of data space from 20 dimensions to 8 dimensions and enhance the fastness of classification process.

Chapter 1

Introduction

The problem of skin cancer is on increase globally and diagnosis of this disease can be done by different techniques. One of these recent techniques is using *Computer Aided Diagnosis* (CAD) systems. In this thesis we used PCA method with NN classifier on assumption to enhance the diagnosis of skin cancer.

1.1 Medical Problem Background

Our body organs are made up of tissues which are ensemble of cells. Cells in their cycles grow, devide and die where the balance between their gowing and daying keeps our bodies functioning normaly. When the cells start to grow without regard to this normal balance, a smal lump of cells called a growth or tumor will be formed (National Cancer Institute, 2011). This growth can be benign (not cancer) or malignant (cancer). Benign growths are not as harmful as malignant growths.

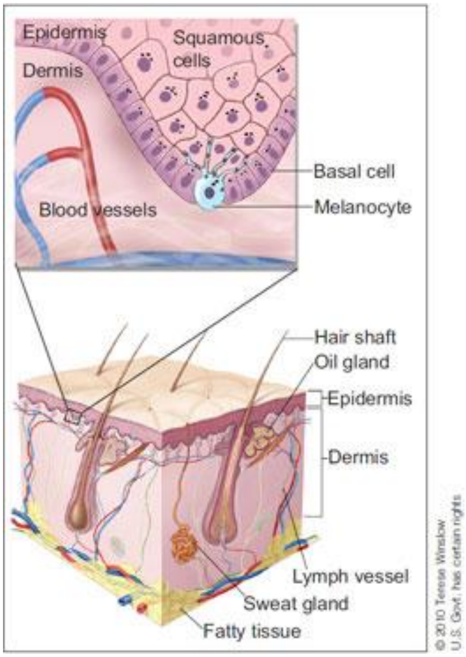


Figure 1. The layers of the skin (National Cancer Institute, 2011)

The outer layer of skin, the epidermis, is made up of different types of cells, as shown in Figure 1. Each type of cancer is distinguished by the cells in which it originates, so skin cancers are classified by the types of epidermal cells involved:

- *Basal Cell Carcinoma* (BCC) develops from abnormal growth of the cells in the lowest layer of the epidermis and is the most common type of skin cancer.
- *Squamous Cell Carcinoma* (SCC) involves changes in the squamous cells, found in the middle layer of the epidermis. BCC and SCC are the most common *non-melanoma* skin cancers.

- *Melanoma* occurs in the *melanocytes* (cells that produce pigment) and is less common than SCC or BCC, but more dangerous. It is the leading cause of death from skin disease.

Skin cancers are sometimes classified as either *melanoma* or *non-melanoma*. (Kantor and Jacksonville, 2009).

In automated diagnosis of skin lesions, feature design is based on the so-called *ABCD rule* of dermatology. *ABCD rule,* which constitutes the basis for *malignant melanoma* diagnosis by a dermatologist, represents:

- Asymmetry: where the shape of one half does not match the other half.
- Border that is irregular: the edges are often ragged, notched, or blurred in outline. The pigment may spread into the surrounding skin.
- Color that is uneven: Shades of black, brown, and tan may be present. Areas of white, gray, red, pink, or blue may also be seen.
- Diameter: There is a change in size, usually an increase. *Melanoma*s can be tiny, but most are larger than the size of a pea (larger than 6 millimeters or about 1/4 inch), (Nachbar, et al., 1994).

*Melanoma*s can vary greatly in how they look. Many show all of the ABCD features. However, some may show changes or abnormal areas in only one or two of the ABCD features.

More information and symptoms about almost all types and subtypes of *melanoma* and *non-melanoma* skin cancers are presented in Appendix A.

1.2 How automated skin cancer recognition works?

This section presents the most common steps used by many researches work on the same field of the proposed system such as (Zagrouba and Barhoumi, 2004, Maglogiannis, et al., 2006, Lau and Al-Jumaily 2009, Salah et al., 2011, Fassihi, et al. 2011). Figure 2 illustrates these steps.

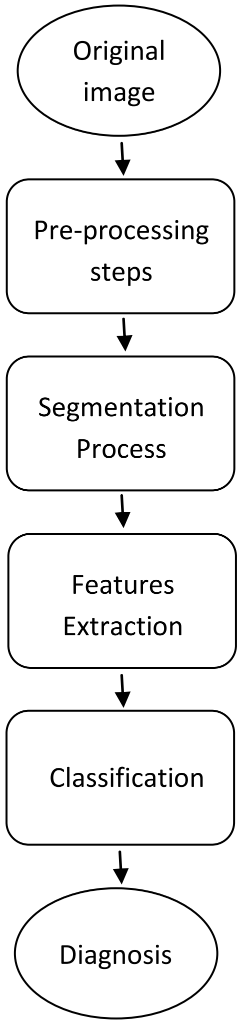


Figure 2. Skin cancer recognition common steps

Pre-processing steps:

For most practical applications, the original input images are typically preprocessed to enhance input images and simplify subsequent operations without losing relevant information. In this stage of system many steps could be performed such as: noise reduction, hair removal, Artifact removal and removing unnecessary features such as light flecks.

The need to use a specific preprocessing step varies depending on the nature of the images used by the application.

Image Segmentation and object detection:

The image is segmented into individual regions to perform a separate analysis of each region. *Region of Interest* (ROI) is detected and localized. So, in result, the skin lesion is isolated from the skin and image background.

Features Extraction:

When the input image data is too large to be processed and it is suspected to be notoriously redundant (much data, but not much information) then the input data will be transformed into a reduced representation set of features (also named features vector). Transforming the high dimensional space of the input data into a lower dimensional space of the set of relevant features is called features extraction.

Feature vectors are computed from the object images. The features are numeric parameters that have to characterize the object.

Briefly, the aim of this stage is to find useful features that are fast to compute, and yet that also preserve useful discriminatory information enabling skin lesion to be distinguished and classified.

Classification:

A feature vector is converted into a decision using a *Machine Learning* (ML) classifier. Classification stage assigns a category to the object based on the feature vector, the input of the classifier, which is provided during feature extraction (Bishop, 2006).

1.3 Problem Scope

There are many types of malignant skin lesions. The main problem of the proposed research is to classify malignant *melanoma* and *non-melanoma* types. This thesis used the digital images of different types of skin cancer (*melanoma* and *non-*

melanoma) from previous research (Salah, et al., 2011) in which 58 digital images of malignant skin lesions were studied, 39 of them were *melanoma* and 19 were *non-melanoma* types. A sample of images used in that research is illustrated in Figure 3.



Figure 3. Types of skin cancer studied in this research (Salah, et al., 2011)

Table 1 shows these types and number of images for every type.

Table 1. Number and types of used skin cancer images

Cancer type	Number of Images
Melanoma	
Superficial Spreading Melanoma	16
Nodular Melanoma	9
Lintigo Maligna Melanoma	10
Acral Lentiginous Melanoma	4
Non-melanoma	
Basal Cell Carcinoma	10
Squamous Cell Carcinoma	7
Other types of non-melanoma	
Sebaceous Gland Carcinoma	2

1.4 Motivation and Significance of Study

Skin cancer is considered the most common cancer all over the world. Early detection of skin cancer affects mortality and morbidity of this disease, especially *melanoma* skin cancer in which the corner stone of cure is early detection.

A preliminary diagnosis of skin cancer is usually done visually by a trained physician, but the definite diagnosis of skin cancer is done by biopsy and histopathological examination, however, this process takes time ranging from 24 to 48 hrs which is the time needed to prepare the tissue so as to be examined by the pathologist (Kumar et al., 2007). Skin cancer patient especially who is diagnosed and treated early have a much higher survival rate.

Dermatology imaging researchers believe that diagnosis of skin tumors can be automated based on certain physical features and color information that are characteristic of the different categories of skin cancer.

Automating the diagnosis of the skin tumors helps in improving the diagnostic accuracy and supports the decision for the suitable next steps especially for new clinician who lacks enough experience. Such a system also is considered a helpful educational tool for teaching physicians and General Practitioners (GPs) to detect and differentiate the types and features of skin lesions.

The proposed research aims mainly to:

- Differentiate *malignant melanoma* among other types of skin cancer.
- Study the efficiency of such researches in case of using PCA to reduce the dimensionality of the data which are used for classification process, especially in terms of fastness.

1.5 Statement of Organization

This thesis is organized as follows:

In chapter 1 a brief outline of the problem and the present work is given.

In chapter 2 a summary of some previous work is provided.

In chapter 3 the adopted images pre-processing and segmentation are presented followed by description of the features extraction.

In chapter 4, using PCA to reduce the dimensionality of data is studied.

In chapter 5 the classification process and its results are provided, and finally the conclusion and future work are presented in chapter 6.

Chapter 2

Previous Work

Since about the last two decades, many researchers like (Ercal, et al., 1994, Maglogiannis, et al., 2005, Blackledge and Dubovitskiy, 2008) have been interested in present systems which are automatically diagnose skin cancer from their digital images. Different types of skin lesions were studied. Image processing techniques and *Artificial Intelligence* (AI) based classifiers were used.

Ercal, et al. (1994) used color images of skin tumors, image processing techniques, and an *Artificial Neural Network* (ANN) to distinguish *melanoma* from three types of benign pigmented tumors: *dysplastic nevi, intradermal nevi* and *seborrheic keratoses*. Features that were expected to distinguish *melanoma* from the three other types of skin tumors (fourteen features) are defined, and then an ANN is used for classification process. They achieved above 80% correct classification of the malignant and benign tumor.

Schmid-Saugeon, et al. (2002) presented a CAD system for pigmented skin lesions. Some solutions for the lesion boundary detection and for the quantification of the degree of symmetry were performed. Feature extraction was restricted in their paper to the quantification of degree of symmetry, even though it was clear that many other features will be necessary for a complete diagnosis system. The symmetry quantification step provided a six dimensional feature vector that can be used to classify pigmented skin lesions as being or malignant.

Zagrouba and Barhoumi (2004) classified skin lesions as malignant or benign using color images of skin lesions. Image processing techniques were used to detect lesion in image and ANN to classify it as malignant or benign. In their research, set of 15 features describing the *ABCD rule* were used. A correct classification rate was measured for every feature alone where the attribute that could decrease the global diagnostic performance was discarded. So, in result, for classification process a set of 9

features was used as input to the classifier. They obtained 79.1% of correct classification of malignant and benign lesions on real skin lesion images.

Maglogiannis, et al. (2005) described a computer-based system for the characterization of pigmented skin lesions in digital images. The first stage included an image acquisition arrangement designed for capturing skin images, under reproducible conditions. The system processed the captured images and performs unsupervised image segmentation on image. Twenty Border and color based features which were extracted from the digital images of skin lesions, were used to construct a classification module for the recognition of *malignant melanoma* versus *dysplastic nevus*. Different methods, AI based method (NNs) and statistical modeling (discriminant analysis), were used in order to find the best classification rules and to compare the results of different approaches to the problem.

Maglogiannis, et al. (2006) discussed intelligent techniques for the segmentation and classification of pigmented skin lesions in dermatological images. A local thresholding algorithm was proposed for skin lesion separation. Border, texture and color based features, were then extracted from the digital images. Extracted features were used to construct a classification module based on *Support Vector Machines* (SVM) for the recognition of *malignant melanoma* versus *dysplastic nevus*. The SVM algorithm was implemented using numerous kernel functions and the best performance was achieved using the exponential radial basis function with sigma=7 (*accuracy*=91.84%, *specificity*=91.87%, *sensitivity*=91.30%).

Blackledge and Dubovitskiy (2008) proposed developing a methodology and implementing applications that were concerned with two key tasks: (i) the partial analysis of an image in terms of its fractal structure and the fractal properties that characterize that structure; (ii) the use of a fuzzy logic engine to classify an object based

on both its Euclidean and fractal geometric properties. The developed system analyzed the structure of a mole, detects cancer identifying features to decide if it is normal or abnormal mole (digital images are used).

Lau and Al-Jumaily (2009) developed a system for early detection of skin cancer by studying and classifying skin cancer images into either benign or *malignant melanoma*. The collected images crossed different image processing procedures to enhance the image properties. Useful information was extracted for the classification process. Tow levels of wavelet decomposition were applied to produce 144 features from the images without clinical knowledge. Two classifiers were studied and compared in the image database that includes dermoscopy photos and digital photos which were: the 3-layers back-propagation NN classifier and auto-associative NN. The overall result was 89.9% for *backpropagation NN* and 80.8% for *auto-associative NN*.

Parolin, et al. (2010) proposed a semi-automated diagnosis of *melanoma* through the analysis of digitized dermatological images. A digital image of the lesion was given to the system, which segmented the lesion automatically using GVF snake technique. A set of eleven features based on the *ABCDE rule* was extracted from lesion pixels. The features extracted from each image were used in a *Bayesian* classifier to identify *melanoma* lesions. Their results indicated that a false negative rate of 1.89% can be achieved, in a total *accuracy rate* of 82.55%.

Przystalski, et al. (2010) proposed decision support system based on semantic analysis method for dermatoscopic images of *malignant melanoma* cases. Thirteen features, 12 colorimetric parameters and one geometric, were extracted form region of interest (ROI) in each image. For classification, ANN and SVMs were used and their performance was evaluated and compared. The success rate in object classification of 93,08% could be achieved by using ANN while it was 97,44% by using linear SVM.

Fassihi, et al. (2011) presented another method to draw distinction between *malignant melanoma* and *benign skin lesion* using the wavelet analysis. Morphological operators were used for segmentation process followed by features extraction using wavelet transform. Eight features, four mean and four variance features, were produced by utilizing the mean and variance of the wavelet decomposition coefficients. These features were the inputs of three layer neural network (NN) which classifies the skin lesion either *malignant melanoma* or benign. Their Results showed 90% ability in distinction between benign and malignant lesions.

Salah, et al. (2011) proposed an automated system that diagnosed six types of cancerous skin lesions. Preprocessing steps were implemented to remove unwanted features impacts from images. By using *thresolding* method, the tumors in digital images were located. Twenty area and color features were calculated to determine the type of skin cancer. For classification process, two techniques were studied and compared: the *hierarchal NN* and *neuro-fuzzy system*. The accuracy rate of the diagnosis of skin cancer by using the hierarchal NN was 90.67% while using neuro-fuzzy system yielded a slightly higher rate of *accuracy* of 91.26% in diagnosis skin cancer type. The *sensitivity* of NN in diagnosing skin cancer was 95%, while the *specificity* was 88%. Skin cancer diagnosis by *neuro-fuzzy system* achieved *sensitivity* of 98% and a *specificity* of 89%.

Chapter 3

Malignant Skin Lesion Localization

and Features Extraction

3.1 Image Pre-processing and Segmentation

Many skin features may have impact on digital images like hair and color, and other impacts such as different size of images, lightness and type of the scanner or digital camera. Pre-processing steps could enhance the images and simplify the next steps such as image segmentation and make it more accurate. The skin cancer image is segmented by the separation of the malignant skin lesion from the healthy skin.

Previous study done by Salah, et al. used the following steps for the image pre-processing:

- Make all images have the same dimensions (Width and Height).
- Removing hair, if any, using *hair removal filter*.
- Color filtration using *BLURE filter*.
- Matching images colors and brightness to be correspondent to images in the collection for high accuracy: using LEVELS tool.

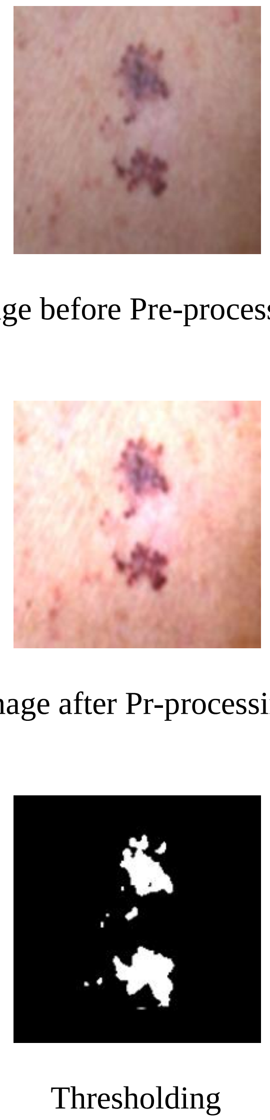


Figure 4. Image Pre-processing and segmentation (Salah, et al., 2011)

And for segmentation process, *Thresholding* method was used. *Thresholding* is based on the idea that the values of pixels belong to the malignant lesion differ from the values of pixels belong to the

background. Color range of images is checked, where red, green and blue color range is between 0 and 250. Pixels that have red color intensity value less than 150, green color intensity value less than 150, and blue color intensity value less than 150 are found to be for the malignant skin lesion.

We adopted the same processed images of malignant skin lesions for the rest of the steps in our research.

3.2 Features Extraction

In this stage, a set of relevant features which can be used to identify the skin cancer type is calculated. These features are extracted by measurements on the image pixels . The features which were used in the present study, were adopted from previous research (Clenek, 1990, Lee and Ercal, 1994, and Salah, et al., 2011), are summarized in Table 2.

Table 2. Skin cancer images features

Feature	Feature description
F1	Irregularity index
F2	Percent asymmetry
F3	Red color variance
F4	Green color variance
F5	Blue color variance
F6	Red relative chromaticity
F7	Green relative chromaticity
F8	Blue relative chromaticity
F9	Spherical color coordinates (L)
F10	Spherical color coordinates (α)
F11	Spherical color coordinates (β)
F12	Color coordinates (L*)
F13	Color coordinates (a*)
F14	Color coordinates (b*)
F15	Red ratio
F16	Green ratio
F17	Blue ratio
F18	Difference in lightness
F19	Difference in chroma
F20	Difference in color

These features are divided into:

- Area Features: Irregularity index and Percent asymmetry features.
- Color Features: all remaining features.

Each feature is calculated using different way and formula (Clenek, 1990, Lee and Ercal, 1994, and Salah, et al., 2011):

a- Area Features:

Images in computer are saved in pixels, pixels are used to find out the area of the tumor as follow:

F1: Irregularity Index:

Irregularity index is calculated using formula (1):

$$I = \frac{P^2}{4\pi A} \qquad (1)$$

Where: P is the perimeter of the tumor in pixels

A is the Area of tumor in pixels

This feature shows the roundness of object, so if the shape of the tumor is a complete circle, the value of P equals "1" which denotes less aggressive tumor.

F2: Percentage Asymmetry

This feature is calculated by dividing the tumor into two halves, overlap the two halves, calculate the non overlapping pixels, and divide this number over the total area in pixels of the tumor as in formula (2):

$$PA = \frac{\Sigma P}{A} \tag{2}$$

Where: P is the number of non overlapping pixels

A is the Area of tumor in pixels

b- Color Features:

The most important features that the classifier depends on is the tumor colors, that are calculated for each pixel using the main colors (R,G,B) as follow:

F3: Variance of RED

Variance of each Red color is calculated using formula (3):

$$Var_{Red} = \frac{\sqrt[2]{\left([\textstyle\sum_{n}^{1} Red_{Tumor,i} - Red_{Tumor,avg} \right)]^2}}{N} \tag{3}$$

$Red_{Tumor,i}$: is the RED value in the i^{th} tumor pixel

$Red_{Tumor,avg}$: is the average of RED color in tumor in pixel

N: total RED pixels in tumor

F4: Variance of GREEN

Variance of each Green color is calculated using formula (4):

$$Var_{Green} = \frac{\sqrt[2]{\left(\left[\sum_{n}^{1} Green_{Tumor,i} - Green_{Tumor,avg}\right)\right]^2}}{N} \quad (4)$$

Where:

Green $_{Tumor,i}$: is the GREEN value in the i th tumor pixel

Green $_{Tumor,avg}$: is the average of GREEN color in tumor in pixel

N: total GREEN pixels in tumor

F5: Variance of BLUE

Variance of each Blue color is calculated using formula (5):

$$Var_{Blue} = \frac{\sqrt[2]{\left(\left[\sum_{n}^{1} Blue_{Tumor,i} - Blue_{Tumor,avg}\right)\right]^2}}{N} \quad (5)$$

Where:

Blue $_{Tumor,i}$: is the BLUE value in the i th tumor pixel

Blue $_{Tumor,avg}$: is the average of BLUE color in tumor in pixel

N: total BLUE pixels in tumor

F6: Relative Chromaticity of RED

Relative Chromaticity of Red color is calculated using formula (6):

$$Red_{chroma,red} = \frac{R_{Tumor}}{R_{Tumor} + G_{Tumor} + B_{Tumor}} - \frac{R_{Skin}}{R_{Skin} + G_{Skin} + B_{Skin}} \quad (6)$$

Where: R_{Tumor} is the total RED values in the tumor

R_{Skin} is the total RED values in the skin

and so on.

F7: Relative Chromaticity of GREEN

Relative Chromaticity of Green color is calculated using formula (7):

$$Green_{chroma,green} = \frac{G_{Tumor}}{R_{Tumor} + G_{Tumor} + B_{Tumor}} - \frac{G_{Skin}}{R_{Skin} + G_{Skin} + B_{Skin}} \quad (7)$$

Where: G_{Tumor} : is the total GREEN values in the tumor

G_{Skin} : is the total GREEN values in the skin

F8: Relative Chromaticity of BLUE

Relative Chromaticity of Blue color is calculated using formula (8):

$$Blue_{chroma,blue} = \frac{B_{Tumor}}{R_{Tumor} + G_{Tumor} + B_{Tumor}} - \frac{B_{Skin}}{R_{Skin} + G_{Skin} + B_{Skin}} \quad (8)$$

Where: B_{Tumor} : is the total BLUE values in the tumor

B_{Skin} : is the total BLUE values in the skin

F9: Spherical Color Coordinates (L)

Spherical color coordinates (L) is calculated using formula (9):

$$L = \sqrt{R^2 + G^2 + B^2} \quad (9)$$

F10: Spherical Color Coordinates (α)

Spherical color coordinates (α) is calculated using formula (10):

$$\alpha = cos^{-1}\left(\frac{B}{L}\right) \tag{10}$$

F11: Spherical Color Coordinates (β)

Spherical color coordinates (β) is calculated using formula (11):

$$\beta = cos^{-1}\left(\frac{R}{L(\sin(\alpha))}\right) \tag{11}$$

Where R Average RED values in the tumor

G Average GREEN values in the tumor

B Average BLUE values in the tumor

F12: Color Coordinates (L*)

Color coordinate (L*) is calculated using formula (12):

$$L* = 116 * \sqrt[3]{\left(\frac{Y}{Y_0}\right)} - 16 \tag{12}$$

F13: Color Coordinates (α*)

Color coordinate (α*) is calculated using formula (13):

$$\alpha * = 500 * \left(\sqrt[3]{\left(\frac{X}{X_0}\right)} - \sqrt[3]{\left(\frac{Y}{Y_0}\right)}\right) \tag{13}$$

F14: Color Coordinates (β*)

Color coordinate (β*) is calculated using formula (14):

$$\beta * = 500 * \left(\sqrt[3]{\left(\frac{Y}{Y_0}\right)} - \sqrt[3]{\left(\frac{Z}{Z_0}\right)} \right) \quad (14)$$

Where X = 2.7690R+1.7518G+1.1300B

Y = 1.0000R+4.5907G+0.0601B

Z = 0.0000R+0.0565G+5.5943B

X0 = 255

Y0 = 255

Z0 = 255

Where 255 is the highest number of R, G, or B in 8 bit color system.

F15: Ratio of tumor color to skin color (R)

Ratio of Red tumor color to skin color is calculated using formula (15):

$$Ratio_{Red} = \frac{R_{Tumor}}{R_{Skin}} \quad (15)$$

Where

R_{tumor} Total of RED values of tumor

R_{skin} Total of RED values of skin

F16: Ratio of tumor color to skin color (G)

Ratio of Green tumor color to skin color is calculated using formula (16):

$$Ratio_{Green} = \frac{G_{Tumor}}{G_{Skin}} \tag{16}$$

Where

G_{Tumor} Total of GREEN values of tumor

G_{Skin} Total of GREEN values of skin

F17: Ratio of tumor color to skin color (B)

Ratio of Blue tumor color to skin color is calculated using formula (17):

$$Ratio_{Blue} = \frac{B_{Tumor}}{B_{Skin}} \tag{17}$$

Where: B_{Tumor} Total of BLUE values of tumor

B_{Skin} Total of BLUE values of skin

F18, F19, F20: Difference in Lightness, Chroma, Hue, and Color

Next formulas show how to calculate difference in Lightness, Chroma, Hue, and Color:

First we must calculate L*, C* and H using formulas (12), (18) and (19):

$$C^* = \sqrt[2]{(a^*)^2 + (b^*)^2} \tag{18}$$

(19)

$$H = tan^{-1}\left(\frac{a^*}{b^*}\right)$$

Where a*, b* are calculated before in formula (13) and formula (14).

F18: Difference in Lightness

Difference in Lightness is calculated using formula (20):

$$\Delta L^* = L^*_{tumor} - L^*_{skin} \qquad (20)$$

F19: Difference in Chroma

Difference in Chroma is calculated using formula (21):

$$\Delta C^* = C^*_{tumor} - C^*_{skin} \qquad (21)$$

F20: Difference in Hue

Difference in Hue is calculated using formula (22):

$$\Delta Clr = \sqrt[2]{\Delta L^{*2} + \Delta a^{*2} + \Delta b^{*2}} \qquad (22)$$

3.3 General Statistics

After establishing the features for the study images sample, related data were analyzed statistically. Nineteen images representing four types of *malignant melanoma*, and nineteen images representing the *non-melanoma* skin cancer were taken and their features values were compared. Figure 5 through Figure 26 illustrate related results.

As shown in Figure 5 and Figure 6 which describe area features, irregularity index and border asymmetry, both *melanoma* and *non-melanoma* skin cancers are close together, and have many intersection points between them which means that the deferentiation between the types of skin cancer based on area features are not dependable.

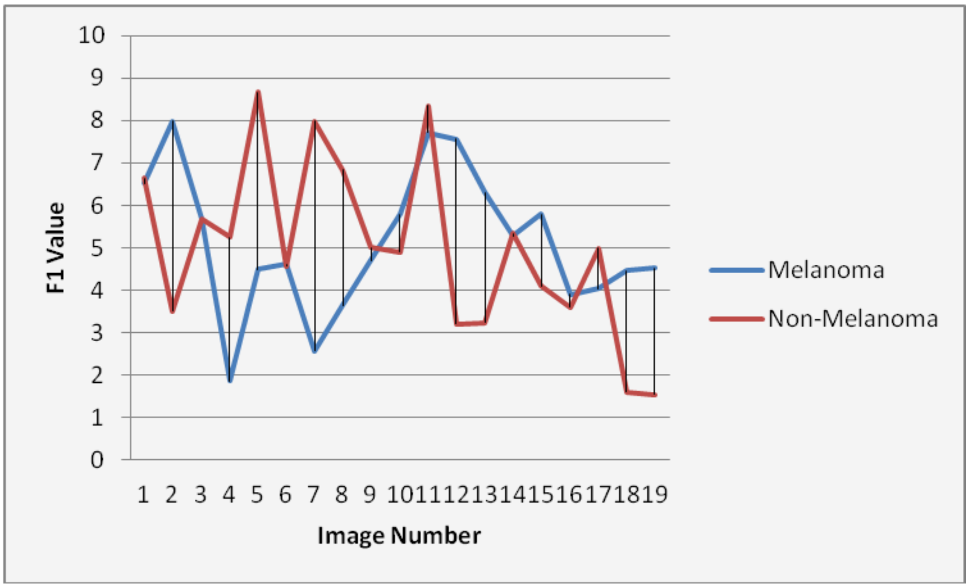


Figure 5. F1 Irregularity Index

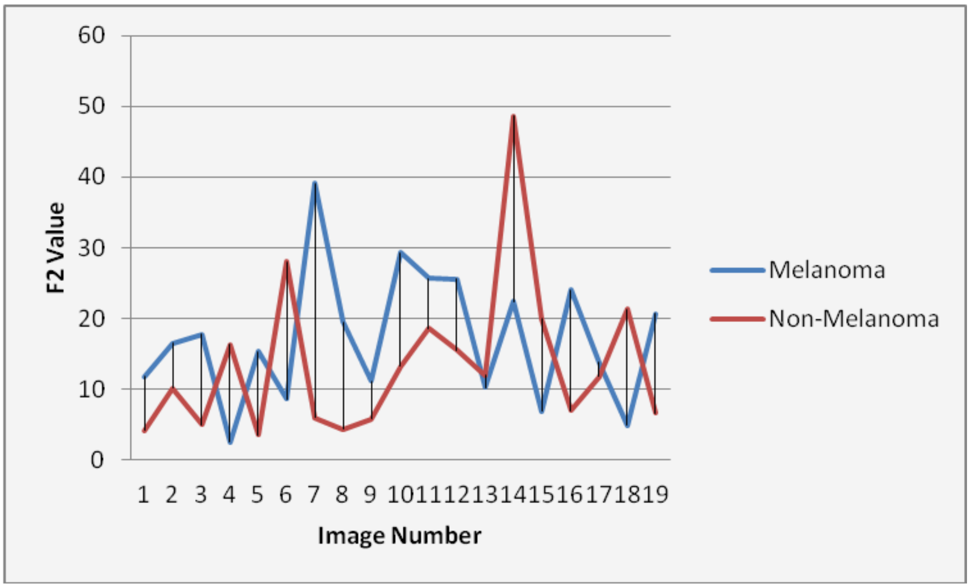


Figure 6. F2 Percent Asymmetry

In general, the color feature is the most dependable feature in recognizing cancer, especially the red color, and we can see in Figure 7 that there is a difference between *melanoma* and *non-melanoma* in the red variance feature, so we can depend on this feature to recognize the type of the skin cancer. Absence of intersection points means that this feature is more dependable to distinguish between *malignant melanoma* and *non-melanoma* tumors.

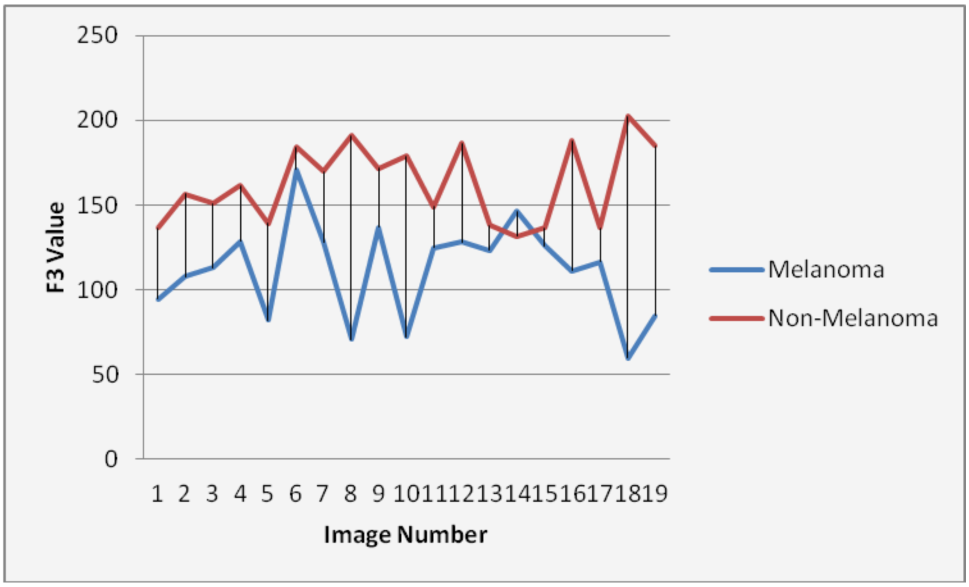


Figure 7. F3 Red Color Variance

Figure 8 shows green color variance feature, there is a similarity between two types of skin cancer. While Figure 9 shows blue color variance feature, which can be used to recognize the type of skin cancer, but there are intersection points between the two types of skin cancer.

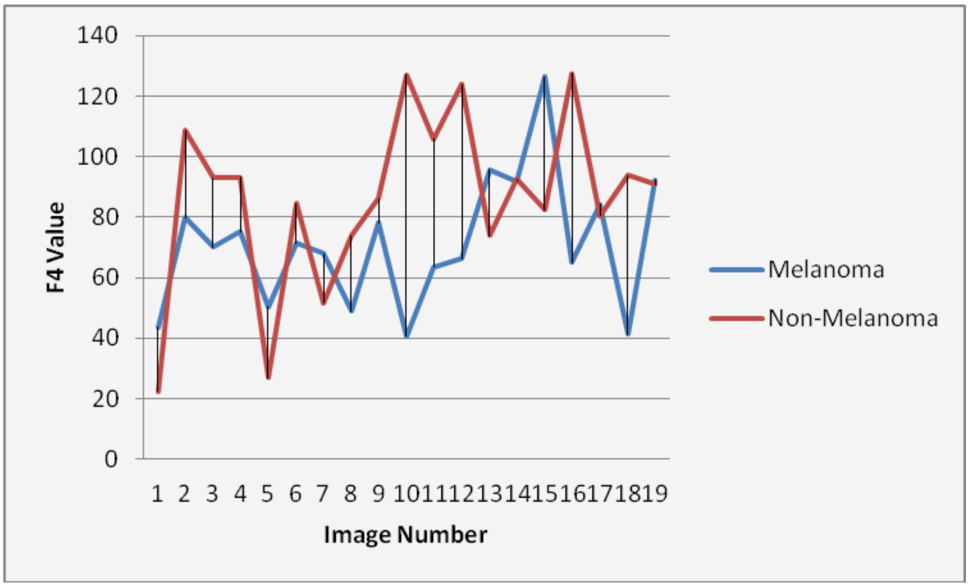


Figure 8. F4 Green Color Variance

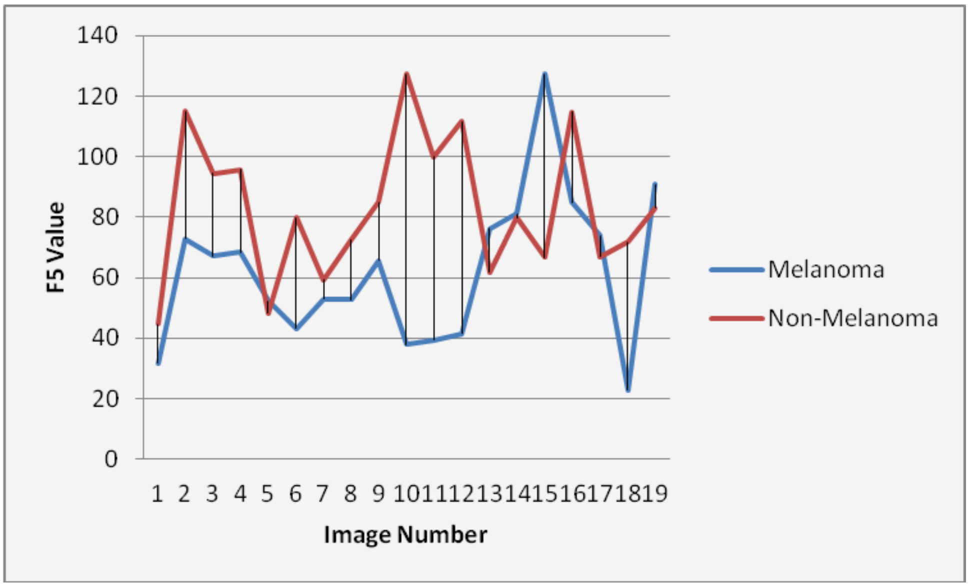


Figure 9. F5 Blue Color Variance

Red relative chromaticity feature shown in Figure 10. *Non-melanoma* skin cancer has greater values than the values of *melanoma*. But green relative chromaticity feature for *non-melanoma* skin cancer has lower values than *melanoma* skin cancer has, as shown in Figure 11. While blue relative

chromaticity feature has too many intersection points between *melanoma* and *non-melanoma* skin cancers as show in Figure 12.

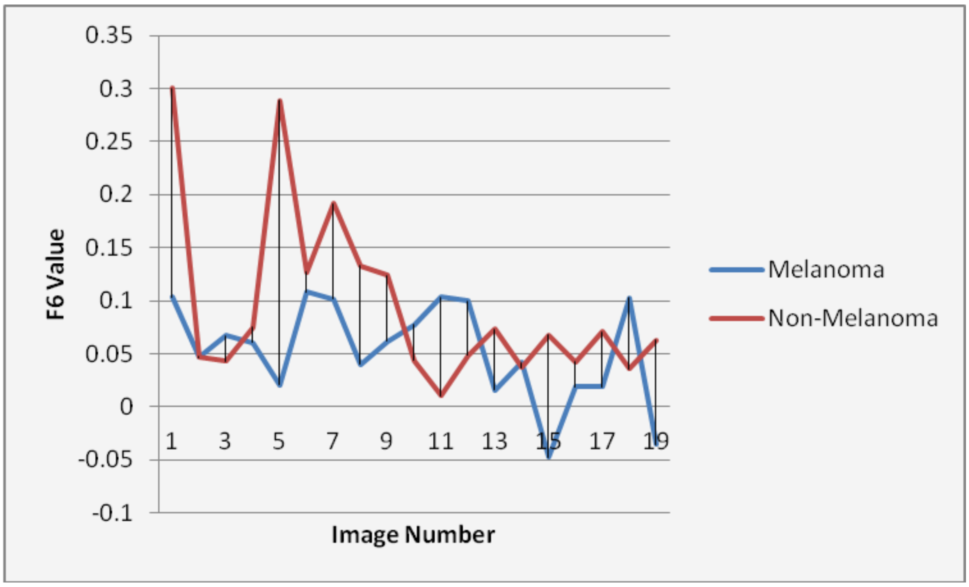


Figure 10. F6 Red Relative Chromaticity

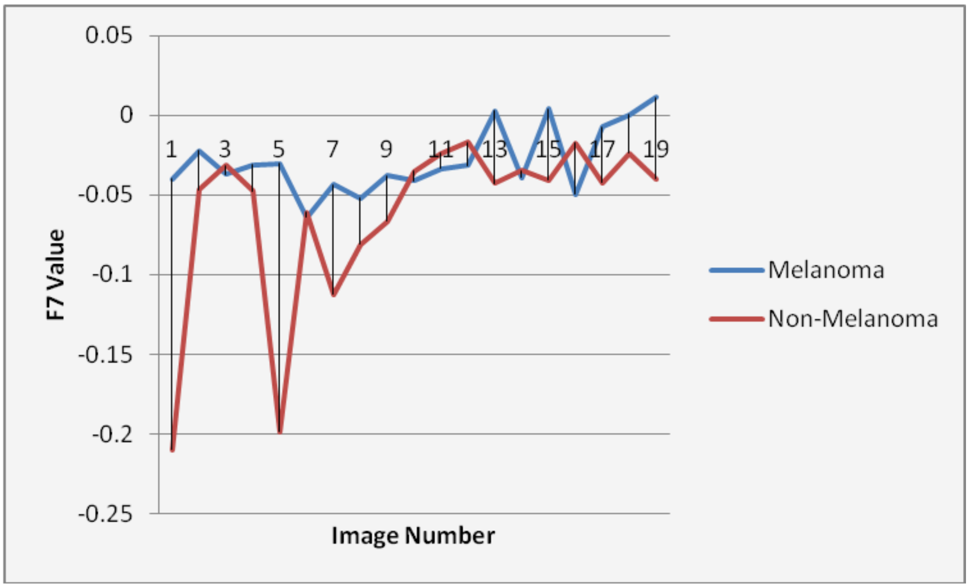


Figure 11. F7 Green Relative Chromaticity

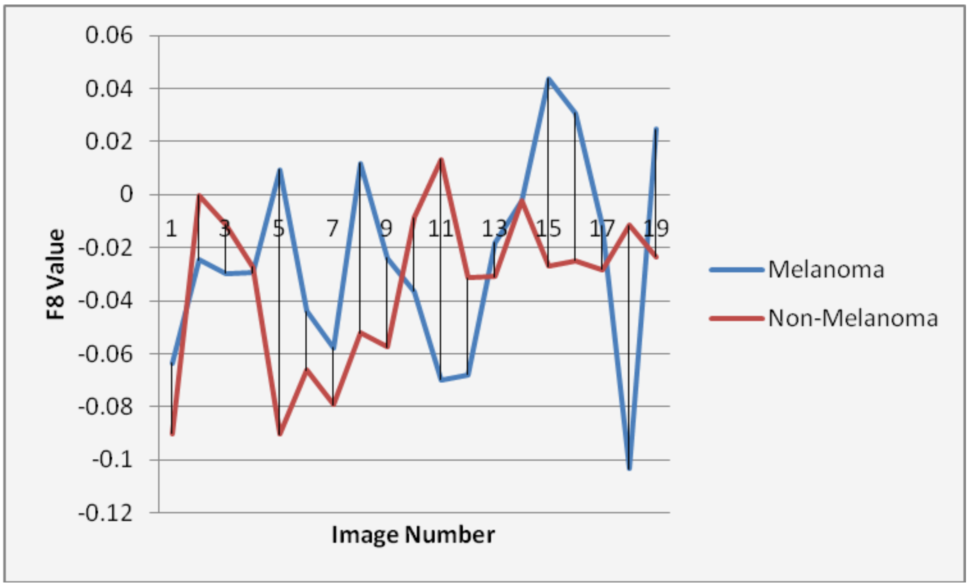


Figure 12. F8 Blue Relative Chromaticity

F9 is another example of dependable feature: For spherical color coordinates L, as shown in Figure 13, it is seen that the value of L is different between the two main skin cancer types (*melanoma* and *non-melanoma*).

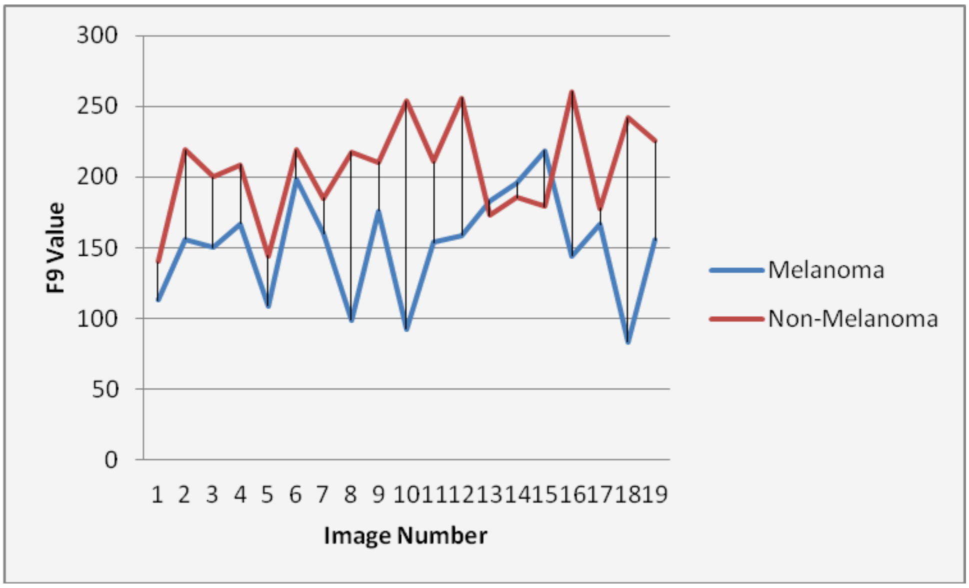


Figure 13. F9 Spherical color coordinates L

Spherical color coordinates α feature has many intersect values between *melanoma* and *non-melanoma* skin cancers, this feature will make skin cancer recognition weak, and this is shown in Figure 14.

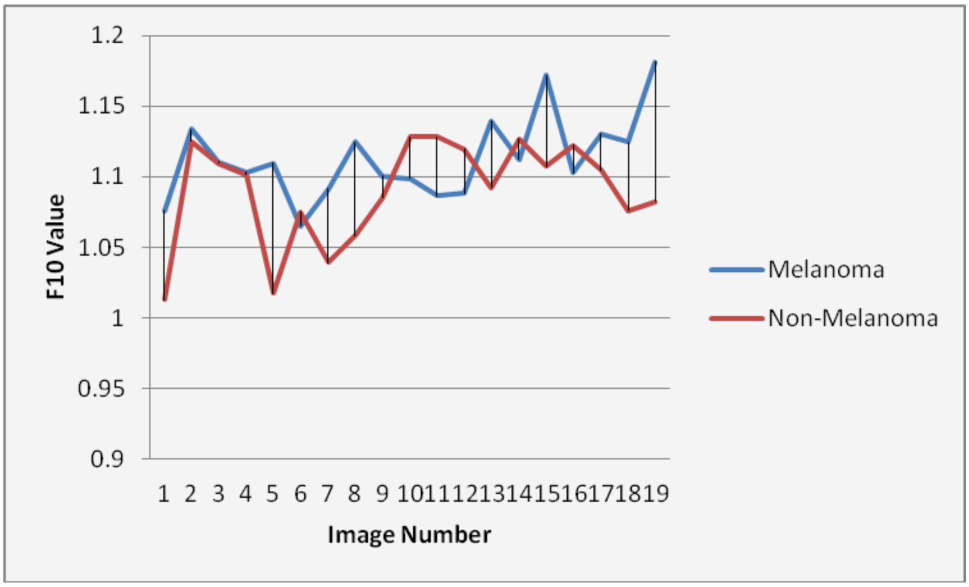


Figure 14. F10 Spherical color coordinates α

Figures 15, 16, 17, 18 show spherical color coordinates β, color coordinates L*, color coordinates a*, and color coordinates b* features respectively. They are weak features to depend on, since they have similar values between *melanoma* and *non-melanoma* skin cancer.

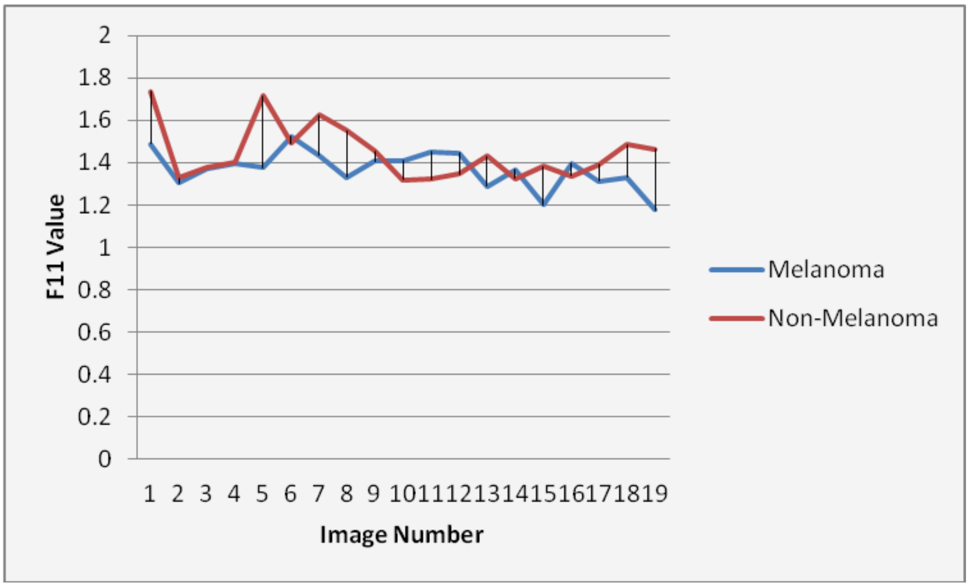


Figure 15. F11 Spherical color coordinates β

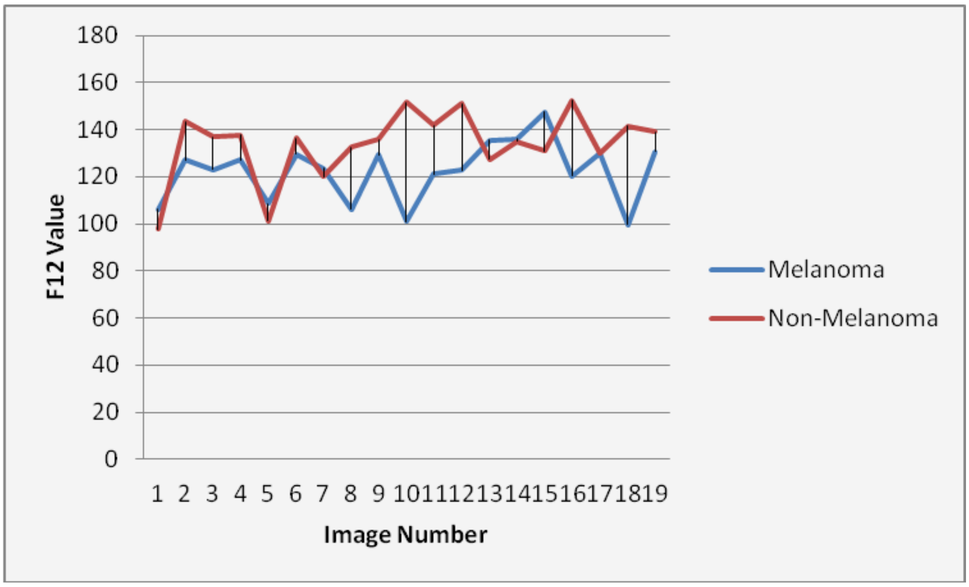


Figure 16. F12 Color Coordinates L*

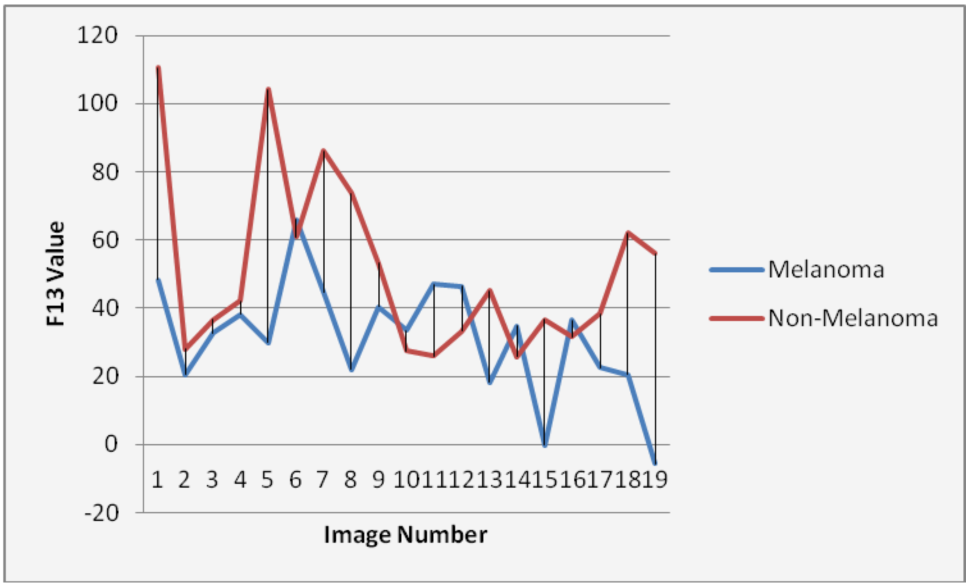


Figure 17. F13 Color Coordinates a*

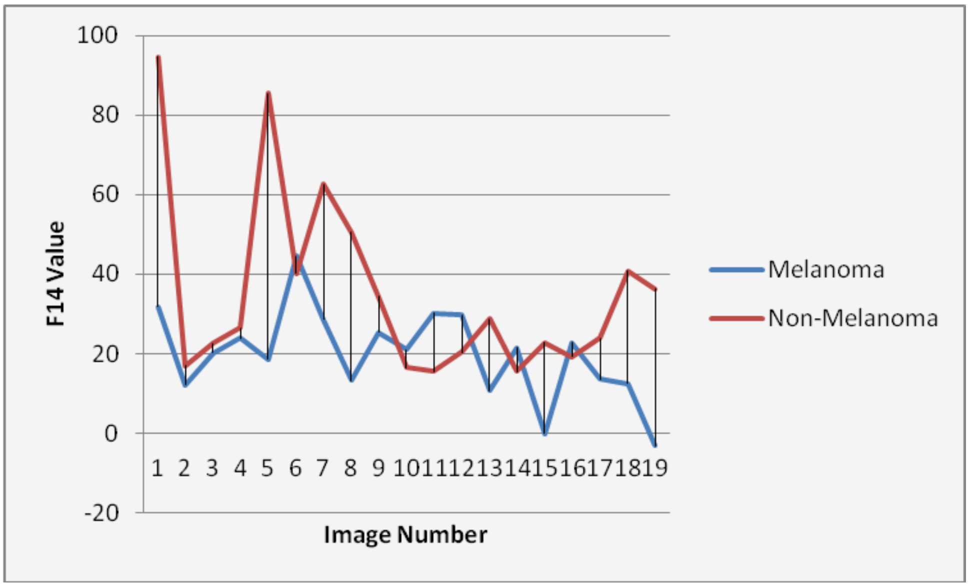


Figure 18. F14 Color Coordinates b*

Figures 19, 23 show red ratio and difference in chroma feature, as we can see no intersection points between *melanoma* and *non-melanoma* skin cancer, they are very powerful features to depend on when recognizing images.

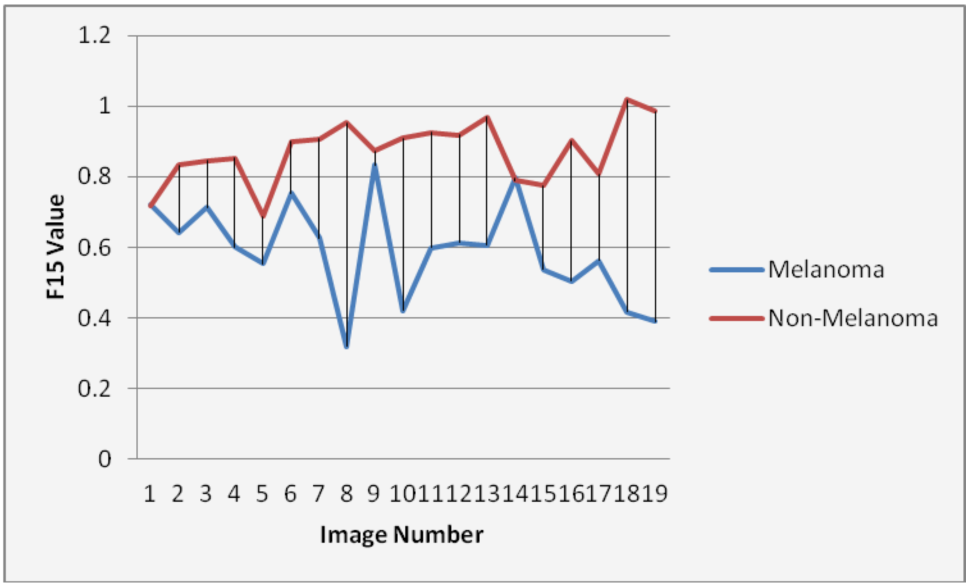


Figure 19. F15 Ratio Red

Figures 20, 21, 22, 24 show green ratio, blue ratio, difference in lightness, and difference in hue features. As we can see many intersection points between *melanoma* and *non-melanoma* skin cancer, so these features are not dependable features, they are weak features to depend on.

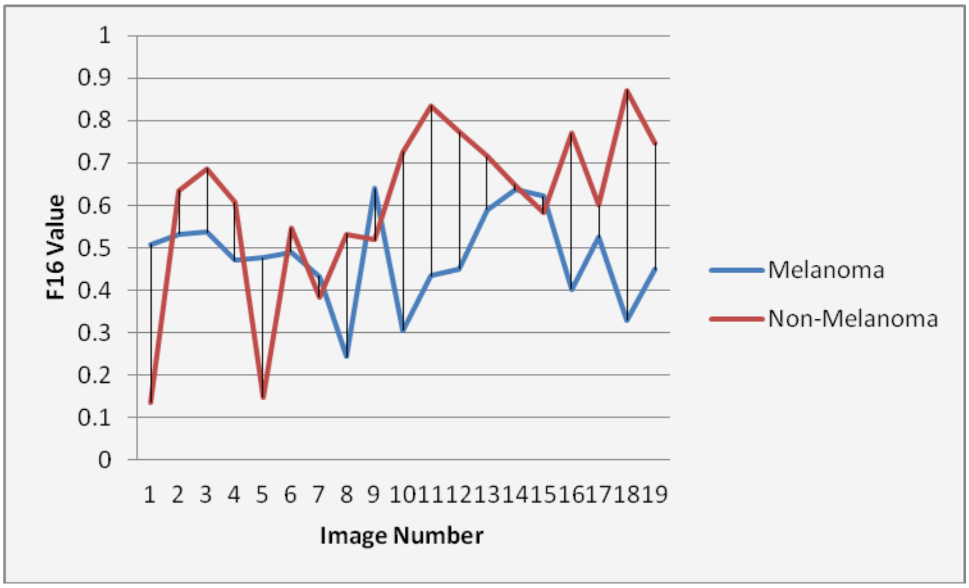


Figure 20. F16 Ratio Green



Figure 21. F17 Ratio Blue

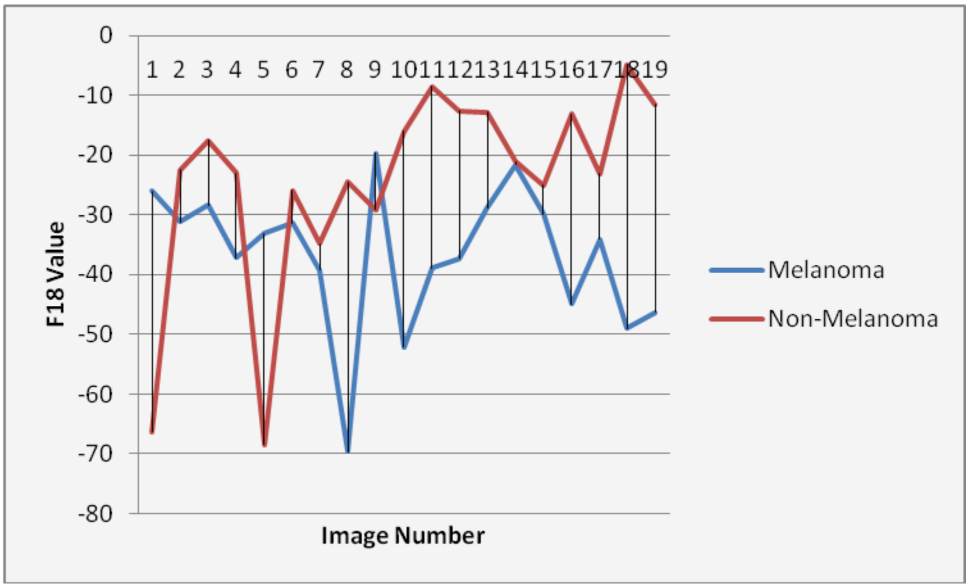


Figure 22. F18 Difference in Lightness

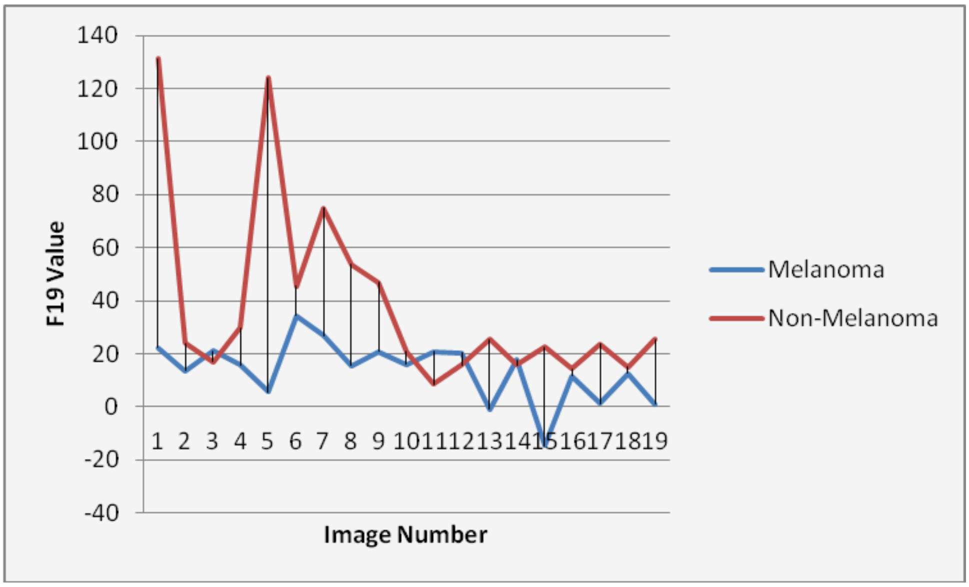


Figure 23. F19 Difference in Chroma

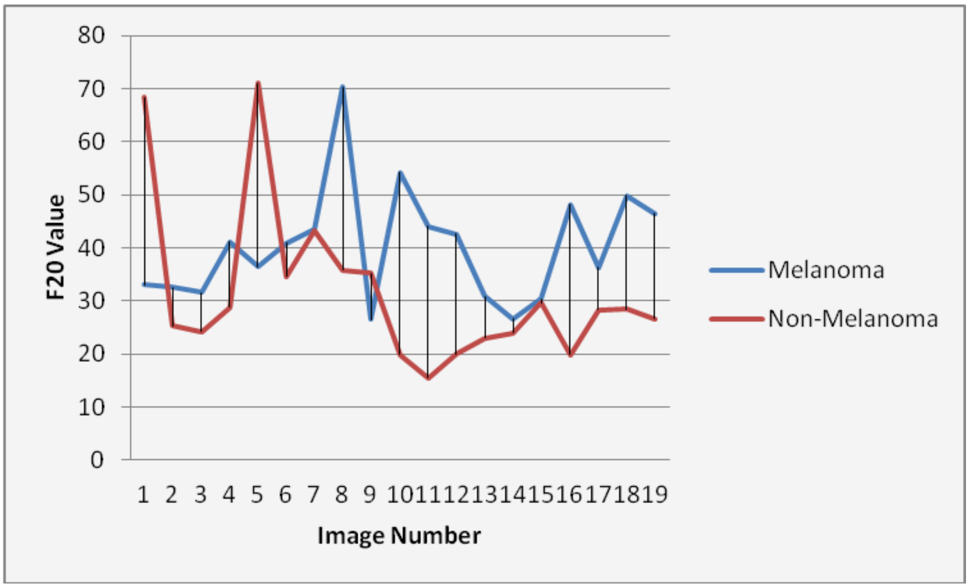


Figure 24. F20 Difference in Hue

As we can see from Figure 25, all features from F1 to F20 of *melanoma* skin cancer have a wide range of values, these gaps occurs due to the differences between each cancer type, such as tumor area, and tumor color. Largest gap was in F20 (difference in hue), while Figure 26 shows all

features values for *non-melanoma* skin cancer, these values are close together which is different from *melanoma* skin cancer.

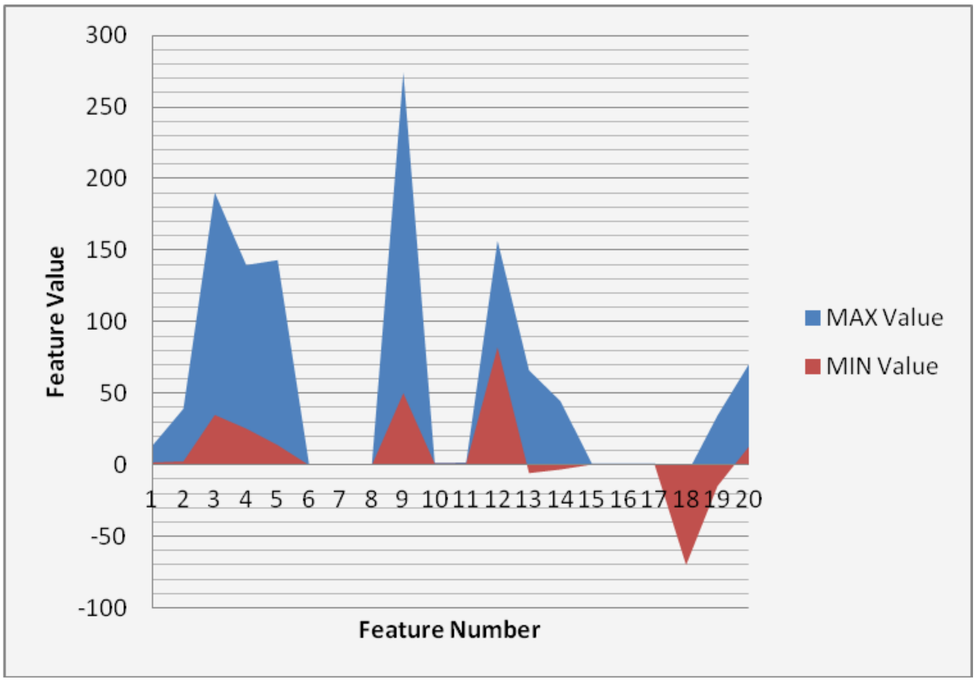


Figure 25. Melanoma F1 to F20 Feature MAX and MIN Values

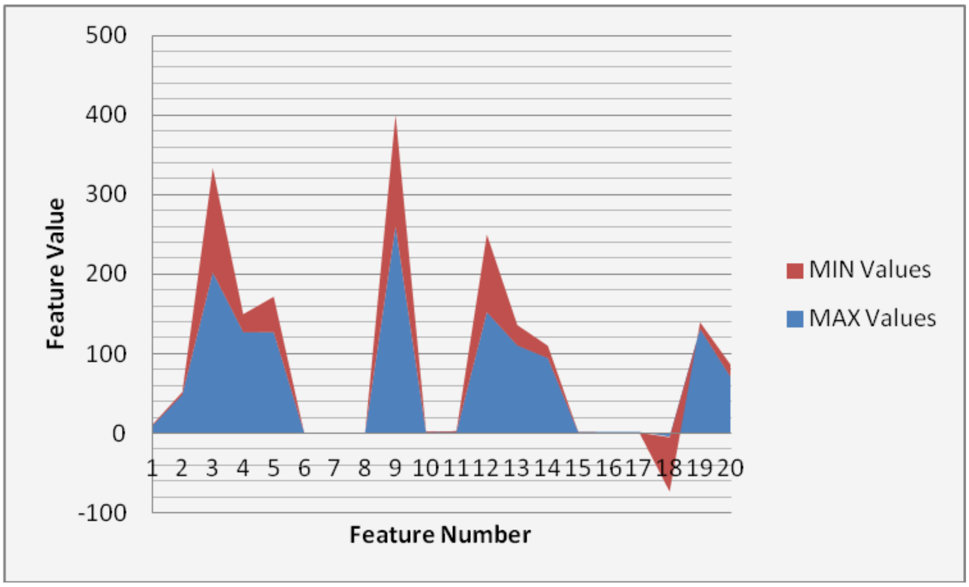


Figure 26. Non-Melanoma F1 to F20 Feature MAX and MIN Values

Chapter 4

Reducing the Dimensions of Data Space

4.1 Introduction

The central idea of PCA is to reduce the dimensionality of a data set consisting of a large number of interrelated variables, while retaining as much as possible of the variation present in the data set. This is achieved by transforming to a new set of variables, the principal components, which are uncorrelated, and which are ordered so that the first few retain most of the variation present in all of the original variables (Jolliffe, 2002).

This transformation is defined in such a way that:

- The first principal component is a single axis in space. When each data point is projected on that axis, the resulting values form a new variable. And the variance of this variable is the maximum among all possible choices of the first axis.
- The second principal component is another axis in space, perpendicular to the first. Projecting the data point on this axis generates another new variable. The variance of this variable is the maximum among all possible choices of this second axis.
- The full set of principal components is as large as the original set of variables. But it is commonplace for the sum of the variances of the first few principal components to exceed 80% of the total variance of the original data (Abdi, 2010).

4.2 Using PCA in the Proposed Research

As mentioned earlier in this research, one of the main goals of this research is to experiment using PCA method to reduce the dimensionality of the data to be used for building NNs to improve classifier performance.

The 20 features which were extracted before is transformed to 20 principal components by PCA method.

The variance carried by each component is illustrated in Table 3 and the total variability of the first five principal components is explained in Figure 27.

Table 3. The percent of the total variability explained by each principal component

Principal component	Percentage of variance (%)
1^{st} comp.	38.4057112842538
2^{nd} comp.	31.5137367300702
3^{rd} comp.	16.2202693262310
4^{th} comp.	5.30522773100576
5^{th} comp.	3.85706261973336
6^{th} comp.	2.54487156429325
7^{th} comp.	1.01648675614382
8^{th} comp.	0.639488479756148
9^{th} comp.	0.158344285966799
10^{th} comp.	0.129983111717699
11^{th} comp.	0.107899977104057
12^{th} comp.	0.0499606191128791
13^{th} comp.	0.0296058755901452
14^{th} comp.	0.0132294502765267
15^{th} comp.	0.00417915655051269
16^{th} comp.	0.00347856254636773
17^{th} comp.	0.000378812348608526
18^{th} comp.	$6.21722924293254 * 10^{-5}$
19^{th} comp.	$2.34793810780393 * 10^{-5}$
20^{th} comp.	$5.62551566131331 * 10^{-9}$

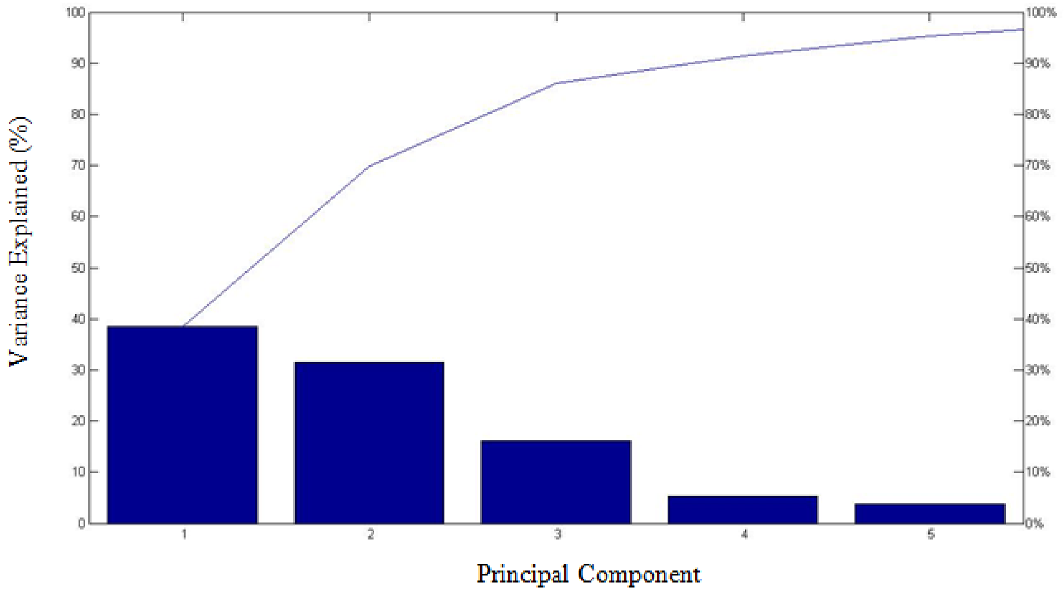


Figure 27. The percent of the total variability explained by the first five principal components

We can notice from Table 3 and Figure 27 that the first three principal components contribute 86.14% to the total variation in the data set and the first eight principal components contribute 99.50% to the total variation in the data set. So the dimensionality of data space is reduced to three dimensions by using the first three components once and to eight dimensions by using the first eight components another time. Classifier performance is experimented with the data represented by each new data space separately.

Chapter 5

Classification using Artificial Neural Networks

5.1 Artificial Neural Network Overview

Neural Networks (NNs) are composed of simple elements operating in parallel. These elements are inspired by biological nervous systems. As in nature, the connections between elements largely determine the network function. You can train a NN to perform a particular function by adjusting the values of the connections (weights) between elements (Beale, et al., 2010).

Typically, NNs are adjusted, or trained, so that a particular input leads to a specific target output. Figure 28 illustrates such a situation. There, the network is adjusted, based on a comparison of the output and the target, until the network output matches the target. Typically, many such input/target pairs are needed to train a network (Beale, et al., 2010). This style of learning is referred to as supervised learning (or learning with a teacher) because we are given the target values.

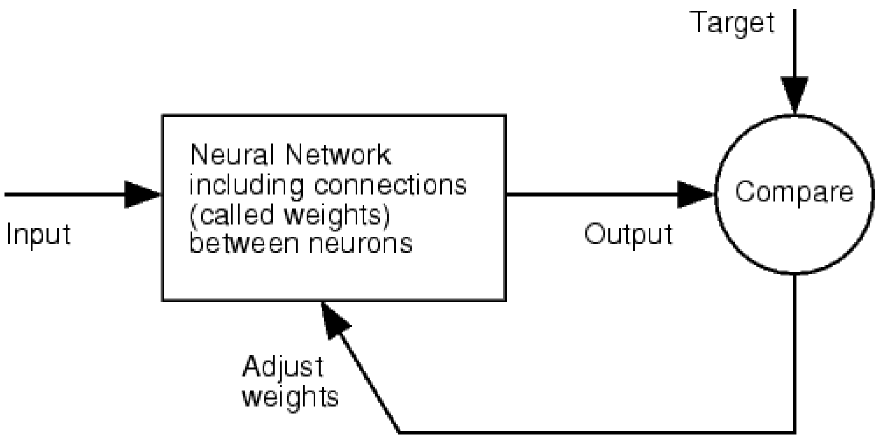


Figure 28. The process of ANN learning (Beale, et al., 2010)

Neural networks have been trained to perform complex functions in various fields, including pattern recognition, identification, classification, speech, vision, and control systems.

5.2 Artificial Neural Network Design

The ANN used in this research is a *Feed-forward NNs* with a two hidden layers. Table 4 shows the main components and their numbers in the used network.

Note: This network will be used many times with different number of inputs as mentioned in the next section, but firstly it will be used with 20 inputs.

The proposed ANN has one output in the third layer where: if output is equal or more than 0.5, the type of malignant skin lesion is *melanoma*. But if it is less than 0.5, it is *non-melanoma* type.

Table 4. The main components of the used ANN

Component	Number of elements
inputs	20
Hidden layers	2
neurons in 1^{st} layer	9
neurons in 2^{nd} layer	5
outputs	1

The overall network function takes the formula (23):

$$y_l = h_3\left(\sum_{k=1}^{N} w_{lk}^{(3)}\ h_2\left(\sum_{j=1}^{M} w_{kj}^{(2)}\ h_1\left(\sum_{i=1}^{D} w_{ji}^{(1)}\ x_i + w_{j0}^{(1)}\right) + w_{k0}^{(2)}\right) + w_{l0}^{(3)}\right) \qquad (23)$$

where *D* is the number of inputs, *M* is the number of neurons in the first layer and *N* is the number of neurons in the second layer. $h_1(.)$, $h_2(.)$ and $h_3(.)$ are activation functions. The superscripts (1) (2) and (3) indicate that the corresponding parameters are in the first, second and third 'layer' of the network respectively. The parameters $w_{lk}^{(3)}$, $w_{kj}^{(2)}$, $w_{ji}^{(1)}$ are weights and $w_{j0}^{(1)}$, $w_{k0}^{(2)}$, $w_{l0}^{(3)}$ are biases. y_l is the output where l =1.

For more illustration, Figure 29 shows the units and connections of the proposed NN:

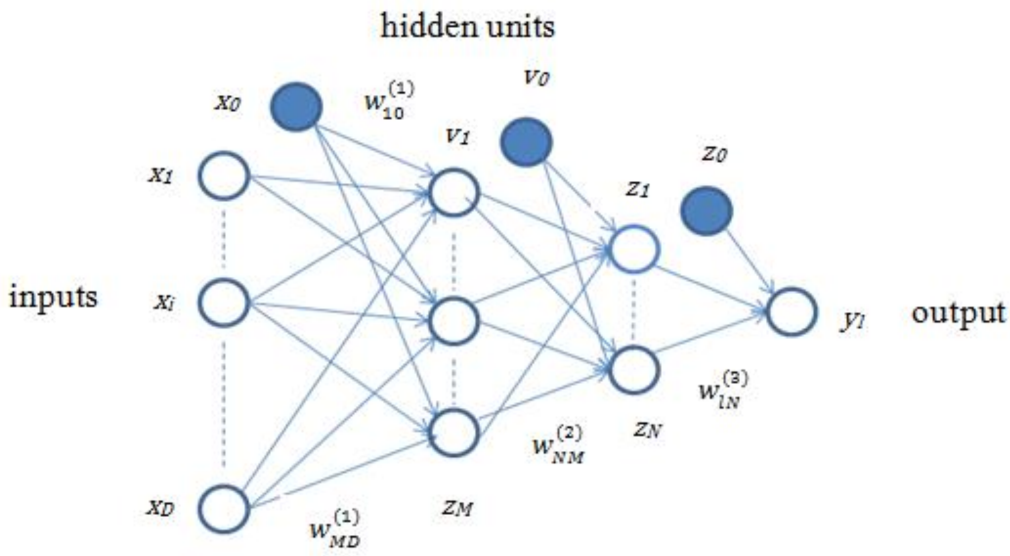


Figure 29. Network diagram for three layers neural network

In Figure 29, the input, hidden, and output variables are represented by nodes, and the weight parameters are represented by links between the nodes, in which the bias parameters are denoted by links coming from additional input and hidden variables x_0, v_0 and z_0. Arrows denote the direction of information flow through the network during forward propagation.

5.3 Classification Results

The previous NN is used to train the 20 extracted features (NN inputs) many times. The same NN is trained with the first three principal components, which contribute more than 80% to the total variation in the data set. Then, it is trained with the all components contribute more than 2% to the total variation in the data set i.e. the first eight components.

The following tables summarize the results of classification process using the proposed NN.

Table 5. Results of training the proposed NN with the 20 features as inputs

No.	Total correct classification	correct classification rate
1	54/58	93.10%
2	53/58	91.38%
3	51/58	87.93%
4	51/58	87.93%
5	53/58	91.38%
6	53/58	91.38%
7	54/58	93.10%
8	53/58	91.38%
9	52/58	89.66%
10	53/58	91.38%

Table 6. Results of training the proposed NN with the first 3 components as inputs

No.	Total correct classification	correct classification rate
1	51/58	87.93%
2	52/58	89.66%
3	50/58	86.21%
4	51/58	87.93%
5	52/58	89.66%
6	50/58	86.21%
7	52/58	89.66%
8	50/58	86.21%
9	49/58	84.48%
10	51/58	87.93%

Table 7. Results of training the proposed NN with the first 8 components as inputs

No.	Total correct classification	correct classification rate
1	51/58	87.93%
2	50/58	86.21%
3	52/58	89.66%
4	51/58	87.93%
5	53/58	91.38%
6	52/58	89.66%
7	54/58	93.10%
8	52/58	89.66%
9	56/58	96.55%
10	52/58	89.66%

Every time the NN is trained, different results are appeared because of the random initialization of the weights parameters.

The average rate of correct classification (*accuracy*) with the 20 features is 90.86%, while it is 87.59% by using the first 3 principal components, and 90.17% by using the first 8 principal components.

Apart from the accuracy indices, the performance of a binary classifier is further evaluated using the *sensitivity* and *specificity* indices. *Sensitivity* and *specificity* rates are calculated according to formula (24) and formula (25):

$$Sensitivity = \frac{TP}{TP+FN} \tag{24}$$

$$Specificity = \frac{TN}{TN+FP} \tag{25}$$

where:

True Positive (TP) indicates *melanoma* types actually classified as *melanoma*, *True Negative* (TN) indicates *non-melanoma* types actually classified as *non-melanoma*, *False Positive* (FP) indicates *melanoma* types actually classified as *non-melanoma* and *False Negative* (FN) indicates *non-melanoma* types actually classified as *melanoma*.

Tables 8 through Table 11 illustrate the performance of NN using the *sensitivity* and *specificity* indices.

Table 8. Performance of the NN with the 20 features as inputs

No.	TP	TN	FP	FN	SENSITIVITY	SPECIFICITY
1	38	16	1	3	92.68%	94.12%
2	36	17	3	2	94.74%	85.00%
3	34	17	5	2	94.44%	77.27%
4	37	14	2	5	88.10%	87.50%
5	35	18	4	1	97.22%	81.82%
6	35	18	4	1	97.22%	81.82%
7	36	18	3	1	97.30%	85.71%
8	36	17	3	2	94.74%	85.00%
9	39	13	0	6	86.67%	100.00%
10	38	15	1	4	90.48%	93.75%

Table 9. Performance of the NN with the first 3 components as inputs

No.	TP	TN	FP	FN	SENSITIVITY	SPECIFICITY
1	35	16	4	3	92.11%	80.00%
2	34	18	5	1	97.14%	78.26%
3	32	18	7	1	96.97%	72.00%
4	37	14	2	5	88.10%	87.50%
5	37	15	2	4	90.24%	88.24%
6	35	15	4	4	89.74%	78.95%
7	38	14	1	5	88.37%	93.33%
8	35	15	4	4	89.74%	78.95%
9	33	16	6	3	91.67%	72.73%
10	34	17	5	2	94.44%	77.27%

Table 10. Performance of the NN with the first 8 components as inputs

No.	TP	TN	FP	FN	SENSITIVITY	SPECIFICITY
1	32	19	7	0	100.00%	73.08%
2	37	13	2	6	86.05%	86.67%
3	38	14	1	5	88.37%	93.33%
4	33	18	6	1	97.06%	75.00%
5	37	16	2	3	92.50%	88.89%
6	33	19	6	0	100.00%	76.00%
7	38	16	1	3	92.68%	94.12%
8	38	14	1	5	88.37%	93.33%
9	38	18	1	1	97.44%	94.74%
10	35	17	4	2	94.59%	80.95%

Table 11.The sensitivity and specificity average of proposed NN

NN input	SENSITIVITY AVERAGE	SPECIFICITY AVERAGE
20 features	93.36%	87.20%
3 components	91.85%	80.72%
8 components	93.71%	85.61%

The time average estimated for classification process is illustrated in Table 12.

Table 12.Time average of classification using proposed NN

NN input	TIME AVERAGE (second)
20 features	1.2051
3 components	1.1434
8 components	1.0576

Chapter 6

Conclusion and Future Work

Conclusion and Future Work

Computer Aided Diagnosis (CAD) systems may serve as diagnostic adjuncts for medical professionals for the confirmation of a diagnosis, as well as for training new dermatologists, plastic surgeons and even General Practitioners, and enhancing the quality of medical care, particularly in areas where a specialized physician or surgeon are not available.

An automated diagnosis system of malignant skin lesions has been presented in this thesis. Skin cancer digital images are processed and classified as a malignant *melanoma* type or *non-melanoma* type.

The images are classified using the extracted 20 features once, the first three principal components, and finally using the first eight components that are calculated by using *Principal Components Analysis* (PCA) method. Feed forward neural network is used as artificial intelligence based classifier.

The first three principal components contribute 86.14% to the total variation in the data set and the first eight principal components contribute 99.50% to the total variation in the data set.

The *accuracy* rate of neural network classifier by using the 20 extracted features as inputs was 90.86% with average time 1.2051 seconds, while it was 87.59% with average time 1.1434 seconds by using the first three components, and 90.17% with average time 1.0576 seconds by using the first eight components. The *sensitivity* rate of neural network by using the 20 extracted features as inputs was 93.36% and the *specificity* was 87.20%, while the *sensitivity* was 91.85% by using the first three components and the *specificity* was 80.72%, and the *sensitivity* was 93.71% by using the first eight components and the *specificity* was 85.61%.

We conclude that the neural network performance indices (*accuracy, sensitivity* and *specificity*) gained using the first eight components are somehow sensible and acceptable. Using the eight components decrease the dimensionality of data space from 20 dimensions to 8 dimensions and enhance the fastness of classification process.

The results of the present thesis are quite promising for the future, so two approaches are proposed to be studied as a future work:

- Enlarging the data set which is used for research.
- Studying the Principal Components Analysis method with statistical based techniques like the *Gaussian distribution* which its main idea is to learn the *mean* and *variance* of each class then compute *the maximum likelihood* for the classification process.

REFERENCES

Abdi, H. and Williams, L.J. (2010). Principal component analysis. **Wiley Interdisciplinary Reviews: Computational Statistics**, 2: 433-459.

Beale, M.H. Hagan, M.T. Demuth, H.B. (2010), **Neural Network Toolbox™ 7 User's Guide,** USA: The MathWorks, Inc.

Bishop, Christopher (2006), **Pattern Recognition and Machine Learning**, (1^{st} ed.). New York: Springer.

Blackledge, J.M. and Dubovitskiy, D.A. (2008), Object Detection and Classifcation with Applications to Skin Cancer Screening. **ISAST transactions on intelligent systems**, 1(2): 1-12.

Brown, D.L. and Borschel, G.H. (2004), **Michigan Manual of Plastic Surgery**, (1^{st} ed.). Philadelphia: Lippincott Williams & Wilkins.

Celenk, Mehmet (1990), A color clustering technique for image segmentation. **Computer Vision, Graphics, and Image Processing**, 52: 145-170.

Ercal, F. Chawla, A. Stoecker, W.V. Lee, H. and Moss, R.H. (1994), Neural Network Diagnosis of Malignant Melanoma From Color Images. **IEEE Transactions on Biomedical Engineering,** 41(9): 837-844.

Fassihi, N. Shanbehzadeh, J. Sarafzadeh, A and Ghasemi, E. (2011), Melanoma Diagnosis by the Use of Wavelet Analysis based on Morphological Operators. **Proceedings of the International MultiConference of Engineers and Computer Scientists**, I: 193-196.

Jolliffe, I. T. (2002), **Principal Component Analysis,** (2^{nd} ed.). New York: Springer.

Kantor, J. and Jacksonville, F. (2009, December 8), Skin cancer – Overview, University of Maryland Medical Center. Retrieved from: **http://www.umm.edu/ency/article/001442.htm**

Kumar, V., Abbas, A. K., Fausto, N. and Mitchell, R. (2007), **Robbins Basic Pathology**, (8^{th} ed.). Philadelphia: Elsevier.

Lau, H.T. and Al-Jumaily, A. (2009), Automatically Early Detection of Skin Cancer: Study Based on Nueral Netwok Classification. **IEEE International Conference of Soft Computing and Pattern Recognition**, 375-380.

Lee, H.C. and Ercal, F. (1994), **Skin cancer diagnosis using hierarchical neural networks and fuzzy logic**. M.S. thesis, University of Missouri, Rolla, MO.

Maglogiannis, I. Pavlopoulos, S. and Koutsouris, D.(2005), An integrated computer supported acquisition, handling, and characterization system for pigmented skin lesions

in dermatological images. **IEEE Transactions on Information Technology in Biomedicine**, 9(1): 86 – 98.

Maglogiannis, I. Zafiropoulos, E. and Kyranoudis, C. (2006), Intelligent Segmentation and Classification of Pigmented Skin Lesions in Dermatological Images, **Lecture Notes in Computer Science**, (3955: 214 – 223), Verlag Berlin Heidelberg: Springer.

Nachbar, F. Stolz, W. Merkle, T. Cognetta, A.B. Vogt, T. Landthaler, M. Bilek, P. Braun-Falco, O. and Plewig, G. (1994), The ABCD rule of dermatoscopy. **Journal of the American Academy of Dermatology,** 30(4): 551-559.

National Cancer Institute (2011, August 1), What You Need To Know About Melanoma and Other Skin Cancers**.** Retrieved from
http://www.cancer.gov/cancertopics/wyntk/skin

Parolin, A. Herzer, E. Jung, C. R. (2010), Semi-Automated Diagnosis of Melanoma Through the Analysis of Dermatological Images. **IEEE Conference on Graphics, Patterns and Images, 71-78.**

Przystalski, K. Nowak, L. Ogorzałek, M. and Surówka, G. (2010), Decision Support System for Skin Cancer Diagnosis. **The Ninth International Symposium on Operations Research and Its Applications**, 406-413.

Salah, B. Alshraideh, M. Beidas, R. and Hayajneh, F. (2011), Skin Cancer Recognition by Using a Neuro-Fuzzy System. **Journal of Cancer Informatics**, 10: 1–11.

Schmid-Saugeona, P. Guillodb, J. Thirana, J. (2002), Towards a computer-aided diagnosis system for pigmented skin lesions. **Computerized Medical Imaging and Graphics,** 27 (2003): 65–78.

Zagrouba, E. and Barhoumi, W. (2004), A Prelimary Approach for the Automated Recognition of Malignant Melanoma. **Image Anal Stereol**, 23: 121-135.

Appendix A Types of Skin Cancer

Melanoma

Melanoma is a form of skin cancer that arises from melanocytes - the cells that produce pigment. Melanoma may begin in association with a mole. Melanocytes produce a pigment called melanin that gives the skin its colour and protects it from sun damage. Darker skin has more melanin and more protection. Melanocytes often cluster together and form moles (nevi). Most moles are benign, but some may go on to become malignant melanomas.

Melanomas are divided into 4 main types, depending on their location, shape and whether they grow outward or downward into the dermis:

Superficial spreading melanoma

- This often begins as a flat dark stain on the skin or appears as a change to a pre-existing mole. It accounts for 2/3 of all melanomas.
- Usually arises from preexisting nevus.
- Affects both genders equally.
- Median age at diagnosis: fifth decade.
- Most common sites: Upper back in men and lower legs in women.
- Irregular, asymmetric borders with color variegation.
- Radial growth phase early; vertical growth phase late.

Nodular melanoma

- Second most common: 15% to 30% of cases.
- Most aggressive type.
- Typically does not arise from pre-existing nevi.

- Men are affected twice as frequently as women.
- Median age at diagnosis is 50 years.
- This is usually unrelated to a pre-existing mole.
- Bluish-black, with uniform, smooth borders.
- It may grow rapidly and spread to the lymph glands quickly.
- Vertical growth phase is a hallmark feature.

Lentigo maligna melanoma (LMM)

- 10% to 15% of cutaneous melanomas.
- LMM is the least aggressive type and the only one clearly associated with sunlight exposure.
- Head, neck, and arms of elderly (sun-exposed areas).
- Women are affected more frequently than men.
- The median age at diagnosis is 70 years.
- Usually greater than 3 cm in diameter, irregular, asymmetric with color variegation; areas of regression may appear hypopigmented.
- Malignant degeneration is characterized by nodular development.

Acral lentiginous melanoma

- About 2% to 8% of melanomas in whites and 35% to 60% of melanomas in blacks, Hispanics, and Asians.
- Presents in palms, soles, and beneath nail plate (subungual). Note: Melanonychia is a linear pigmented streak in the nail, which is often benign and is more common in black and Asian populations. Due to the risk of melanoma, biopsy of suspect lesions should be performed.
- Median age at diagnosis is approximately 60 years.
- Irregular pigmentation is common.

- Large size (>3 cm).
- Majority involve great toe or thumb.
- Long radial growth phase: transition to vertical growth phase occurs with high risk of metastasis.

Basal cell carcinoma

Basal cell carcinomas are usually:

- a single, small, and firm lesion
- dome shaped and flesh-colored.
- raised at the edges.
- pearly or milky white bordered.
- covered with tiny blood vessels - easily seen over the top and around the tumor.
- may look like a pimple that won't heal - ulcerated and bleeding at the centre.

Nodular BCC

Clinically, nodular BCCs comprise 60% of all BCCs. These start off as flat lesions that are best seen when you stretch the skin. Stretching the skin helps you define the edges of the lesion. They slowly grow into small bumps or papules; the most common description of nodular BCC is "pearly papules with a rolled edge." By a "rolled edge," we mean that so many of these basal cells start off being dome-shaped, then as they grow laterally, the middle sinks and there is a little ripple all around the edge of the lesion. And on the surface within there will be little tortuous blood vessels. As the lesions grow, they retain the slightly raised border and, even very early on, when the skin is stretched they become bumps. There is a tendency for them to ulcerate in the middle. As they advance, the middle collapses, so if you leave them long enough, these nodular basal cells present as ulcers. In the past, they were called rodent ulcers because they just continue to eat away. They grow laterally and can infiltrate deeply, which is

the problem. Most nodular BCCs are on the face and, if left untreated, they are particularly destructive around the nose and the eyes. Usually, nodular BCCs are very well-defined, which is important for treatment: you know where they begin and where they end. This distinguishes them from the ones that are ill-defined. The latter are more difficult to treat.

Superficial spreading BCC

Superficial BCCs are thought to be about 15% of BCCs. Because sBCC spreads outward, it was originally thought that it was a multifocal cancer, each developing individually. But it is actually all linked together and part of the same single lesion. Superficial BCC occurs as downgrowths from the undersurface of the epidermis. Superficial BCCs all start off in the epidermis or in the lining of hair follicles, and grow down into the dermis. There are not clumps of basal cell in the dermis, as with other BCCs. Superficial BCCs are erythematous, slightly scaly, well-defined patches, most commonly found on the trunk and the limbs, not usually on the face. They are easily confused with a patch of psoriasis or eczema. Psoriasis usually consists of multiple lesions — the same with eczema— and psoriasis tends to be symmetrical. An sBCC is just a lonely little patch of red, well-defined, slightly scaly skin that may not change very much for months or even years. That, in itself, is a clue. If it is just slowly growing, and the patient does not have other, similar lesions, then it might be an sBCC. The features of sBCC are fairly distinctive upon closer inspection. It too has a very fine beading or fine pearly edge — when the skin is stretched — but there might be some fine scaling and some crusting in the middle. As the sBCCs mature, the middle part tends to smooth out and almost look a little bit scarred. It becomes a little bit atrophic, thinner, smoother and less red. Superficial BCC is a commonly misdiagnosed disease; mostly misdiagnosed as psoriasis, discoid eczema or fungus .

Morpheaform (sclerosing, fibrosing) BCC

The fibrosing or sclerotic BCCs are usually found on the face. They generally present as what looks like a scar, but without any history of trauma. They have a lot of very dense fibrous stroma, which is what makes them look like scars. Fibrosing BCCs have small, scattered basal cells throughout, making them behave more aggressively. It is more difficult to get rid of them. They are usually flat or slightly depressed, can be yellowish in colour and the surface tends to be smooth and shiny. They are often ill-defined at the borders. Unlike nodular BCC, which generally is soft and gelatinous, fibrosing BCCs are firm, so it is not possible to completely curette them. We can scrape off nodular BCCs because they are much softer than the surrounding skin. Fibrosing BCCs are also more diffuse, which is detectable in the pathology. Therefore, the treatment is different for sclerotic BCCs.

Pigmented BCC

Pigmented BCCs behave like nodular BCCs, but they have pigment within them that can either be flecks of brown or black scattered through the lesions, or they sometimes can be completely pigmented. This is very important in the differential diagnosis: pigmented BCCs behave like nodular basal cells, but they can be easily confused with melanomas. Conversely, some melanomas do not have a lot of pigment in them. These amelanotic melanomas — which are scary because the risk of missing them is so great — can be a little bit raised, the skin colour is sometimes red and they have flecks of pigment in them. Pigmented BCCs are not that common, but they do occur .

Adnexal BCC

Uncommon and found in older individuals. Tumors arise from sweat glands, and although they exhibit slow growth, they are locally invasive, with a high incidence of local recurrence.

Fibroepithelioma of Pincus

Fibroepithelioma of Pincus is rare and different from other BCCs. It tends to be a smooth, elevated little nodule that is sometimes pedunculated. It almost looks like a firm skin tag. It is most often seen on the back or the extremities, but can also be found in the groin or on the sole of the foot. As those are not sun-exposed areas, this disease is probably not sun-related.

Squamous cell carcinoma (SCC)

The signs and symptoms of squamous cell carcinoma are reddened, scaly or plaque-like areas that:

- are slightly elevated.
- can develop an ulcer.
- change over time into a hard pale pink to white nodule.
- can develop a scaly crust or a wart-like surface.
- arise from old scars, change colour, begin to bleed, ulcerate, drain, or become painful.
- commonly appear on face, ears, neck, forearms, backs of hands, and legs.

Verrucous SCC

- Slow-growing, exophytic, and less likely to metastasize.

Ulcerative SCC

- Grows rapidly and is locally invasive.
- Ulcerative SCC has very aggressive growth characteristics, raised borders, and central ulceration.
- Less than 50% 5-year survival if spread to lymph nodes in head and neck.

Other Types of Skin Cancer

There are several types of skin cancers which are neither basal, squamous, or melanoma. These are much more rare, however they do still occur and are treatable .

Merkel cell carcinoma

- rare, potentially aggressive form of non-melanoma skin cancer .
- begins in the cells that produce hormones involved in the regulation of the nervous system .
- tumours are found on or just beneath the skin and in the hair follicles.
- Malignant neuroendocrine tumor arising within the dermis from cells of neural crest origin.
- Aggressive, with radial spread, high local recurrence, and regional and systemic metastasis.
- Presents as a purple to red papulonodule or indurated plaque; 50% involve the head and neck, 40% involve the extremities, and 10% involve the trunk.
- Treatment involves wide (up to 3 cm) margins.
- 30-50% survival at 5 years.

Microcystic adnexal carcinoma

- rare low grade tumour of the sweat glands which is unlikely to spread, but can grow locally, becoming quite large.
- Presents as a white to pink papule-plaque primarily on the head and neck.

Sebaceous gland carcinoma

- arises in the sebaceous glands that produces sebum which keeps the skin moist.
- the tumours can be aggressive, and can invade the local and surrounding tissue in several areas.
- can spread in 14-25% of cases to lungs, liver, brain, bone and lymph nodes.
- Yellowish to pink, slowly growing papulonodule on eyelid (resembles chalazion).

Cutaneous T-Cell Lymphoma

- cancer of the lymphatic system that shows a predisposition to attack the skin – normally it attacks the lymphocytes (the blood cells which regulate immunity).
- quite rare, as 90% of lymphomas are B-cell, and in general difficult to treat with a greater chance of recurring.

Kaposi's Sarcoma

- once was extremely rare, only occurring in older people of Russian-Jewish, Italian and African extraction.
- it has had a significant increase of a more virulent form associated with AIDS, is caused by Herpes virus 8 infection, with men affected over twice as much as women.

Atypical Fibroxanthoma

- a rare spindle cell tumour of the skin.
- locally aggressive, it can involve surrounding tissues but is unlikely to spread.
- if it does spread, it may spread to the parotid gland, lymph nodes, and lung.

More medical details could be found in (Brown and Borschel, 2004).

الأتمته السريعة لتشخيص السرطانات الجلدية

إعداد
أمان علي محمد عومان

المشرف
الأستاذ الدكتور محمد بلال الزعبي

المشرف المشارك
الدكتورة بارقة صلاح

ملخـــــــــص

يعد سرطان الجلد أحد أكثر أنواع السرطانات انتشاراً في العالم. تحدث نحو 75% من وفيات سرطان الجلد بسبب الورم الميلانيني الخبيث، في حين أن معدل الشفاء منه ما يقارب 99% إن تم اكتشافه وعلاجه في مرحلة مبكرة .

أنظمة تشخيص السرطان الجلدي المؤتمته تعمل كأدوات مساعدة لتأكيد التشخيص فضلا على أنها تساعد في تدريب وتحسين كفاءة الرعاية الطبية، ولا سيما في الأماكن التي لا يتوفر فيها أخصائيو الأمراض الجلدية .

تقدم هذه الرسالة نظاماً مؤتمتا لتشخيص الأورام السرطانية الجلدية من خلال الصور الفوتوغرافية للأورام. تناولت العديد من الدراسات تقنيات معالجة الصور لتحديد مكان الورم في الصورة، ومن ثم استخلاص سمات معينة لهذا الورم لاستخدامها في مرحلة التصنيف. تم استخدام تقنية الشبكات العصبونية الاصطناعية لتصنيف وتمييز الأورام الميلانينية من أنواع الأورام غير الميلانينية .

ومن أحد أهم سمات هذه الرسالة أنها تدرس فاعلية استخدام تقنية تحليل المركبات الرئيسية للبيانات الممثلة بالسمات المستخلصة مسبقا للتقليل من حجم البيانات المستخدمة في التصنيف وتسريع عملية التصنيف وبالتالي تحسين الأداء العام للنظام. تم تمثيل البيانات بثلاث مركبات رئيسية، و من ثم تم تمثيلها بثمان مركبات رئيسية المحسوبة باستخدام تقنية تحليل المركبات الرئيسية.

كان معدل دقة تصنيف البيانات بعشرين سمة مستخلصة من صور الأورام هو 90,86% بمتوسط وقت 1,2051 ثانية. في حين كان معدل تصنيف البيانات ممثلة بثلاث مركبات رئيسية هو 87,59% و بمتوسط وقت 1,1434 ثانية. و أخيراً كان معدل تصنيف البيانات ممثلة بثمان مركبات رئيسية هو 90,17% و بمتوسط وقت 1,0576 ثانية.

من المؤشرات السابقة لأداء الشبكة العصبونية المستخدمة في عملية التصنيف، خلصنا الى أن تصنيف البيانات في حال تمثيلها بالمركبات الرئيسية الثمان كانت مقبولة عدا على أنها زادت من سرعة تصنيف البيانات و قللت الفضاء البعدي الممثلة فيه البيانات من عشرين بعداً الى ثمانية أبعاد.